ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΣΗΜΕΙΩΣΕΙΣ ΠΙΘΑΝΟΤΗΤΩΝ

Γιάννης Κοντογιάννης, Σταύρος Τουμπής

Ο.Π.Α. 2012

Οδηγίες Χρήσης

Το παρόν ΔΕΝ είναι διδακτικό βιβλίο. Είναι οι σημειώσεις των διαλέξεων και των φροντιστηρίων του μαθήματος «Πιθανότητες» όπως αυτό διδάσκεται στο Τμήμα Πληροφορικής του Οικονομικού Πανεπιστημίου Αθηνών. Το μάθημα διδάσκεται σε ένα εξάμηνο (το δεύτερο), και ό,τι περιέχεται εδώ καλύπτεται σε περίπου 40 δίωρες διαλέξεις (τα φροντιστήρια είναι ενσωματωμένα στις διαλέξεις). Οι σημειώσεις δίνονται ηλεκτρονικά στους φοιτητές.

Η διαμόρφωση της ύλης του μαθήματος, και συνεπώς των σημειώσεων, έχει γίνει λαμβάνοντας υπ' όψιν τις υφιστάμενες γνώσεις των φοιτητών και το γεγονός ότι το μάθημα διδάσκεται στο δεύτερο εξάμηνο ενός τμήματος Πληροφορικής.

Δεν υπάρχει υποκατάστατο της σε βάθος μελέτης ενός καλού βιβλίου για την κατανόηση του αντικειμένου, όπως και δεν υπάρχει υποκατάστατο της φυσικής παρουσίας στο αμφιθέατρο. Συνεπώς, δεν συνιστάται στους φοιτητές, ούτε να χρησιμοποιήσουν τις σημειώσεις σαν βιβλίο, ούτε ως υποκατάστατο της παρακολούθησης. Αντίθετα, ο σκοπός τους είναι:

1. Να διευκολύνουν την παρακολούθηση των φοιτητών, μια που δεν θα χρειάζεται να αντιγράφουν ό,τι γράφεται στον πίνακα, συνήθως υπό δυσμενείς συνθήκες.

2. Να διευκολύνουν την μελέτη της ύλης που διδάχτηκε σε συνδυασμό με τα διδακτικά βιβλία.

3. Να βοηθούν τους φοιτητές που δεν παρακολούθησαν κάποιες διαλέξεις να μείνουν σε επαφή με το μάθημα.

Επίσης οι σημειώσεις ίσως βοηθήσουν ώστε όσοι φοιτητές βρίσκονται στο αμφιθέατρο να μην είναι εκεί απλώς για να «πάρουν τις σημειώσεις».

Το παρόν κείμενο είναι υπό διαρκή εξέλιξη. Θα εκτιμήσουμε ιδιαιτέρως οποιαδήποτε πρόταση για τη βελτίωσή του όπως και τυχόν παρατηρήσεις σχετικά με ορθογραφικά/τυπογραφικά σφάλματα, λάθη στις ασκήσεις και κάθε φύσεως πρόβλημα. Θα είναι χαρά μας επίσης αν το χρησιμοποιήσουν και άλλοι διδάσκοντες.

Γιάννης Κοντογιάννης, `yiannis@aueb.gr`
Σταύρος Τουμπής, `toumpis@aueb.gr`

Περιεχόμενα

Κεφάλαιο 1

Μέτρο Πιθανότητας

1.1 Σύνολα

Ένα μεγάλο μέρος του μαθηματικού λεξιλογίου της Θεωρίας Πιθανοτήτων χρησιμοποιεί βασικά στοιχεία της Θεωρίας Συνόλων. Ξεκινάμε υπενθυμίζοντας μερικούς ορισμούς.

Ορισμός 1.1. (Πράξεις μεταξύ συνόλων)

1. *Ένα σύνολο είναι μια συλλογή στοιχείων. Π.χ., τα* $A = \{-1, +1\}$, $B = \{3, 5, 9\}$, $\mathbb{Z} = \{\dots, -1, 0, 1, 2, \dots\}$ = *οι ακέραιοι αριθμοί*, $\mathbb{R}$ = *οι πραγματικοί αριθμοί*, $C = \{A, B, 5, \{5\}, \mathbb{R}\}$, $D = \{$*Παναγιώτης, Θανάσης*$\}$ *είναι όλα σύνολα.*

2. *Όταν κάποιο στοιχείο* α *ανήκει σε κάποιο σύνολο* A, *γράφουμε* $\alpha \in A$. *Αν το* α *δεν ανήκει στο* A, *γράφουμε* $\alpha \notin A$. *Π.χ., πιο πάνω έχουμε* $3 \in B$, *αλλά* $0 \notin A$.

3. *Το* A *είναι υποσύνολο του* B *αν κάθε στοιχείο του* A *ανήκει και στο* B, *οπότε γράφουμε* $A \subseteq B$.

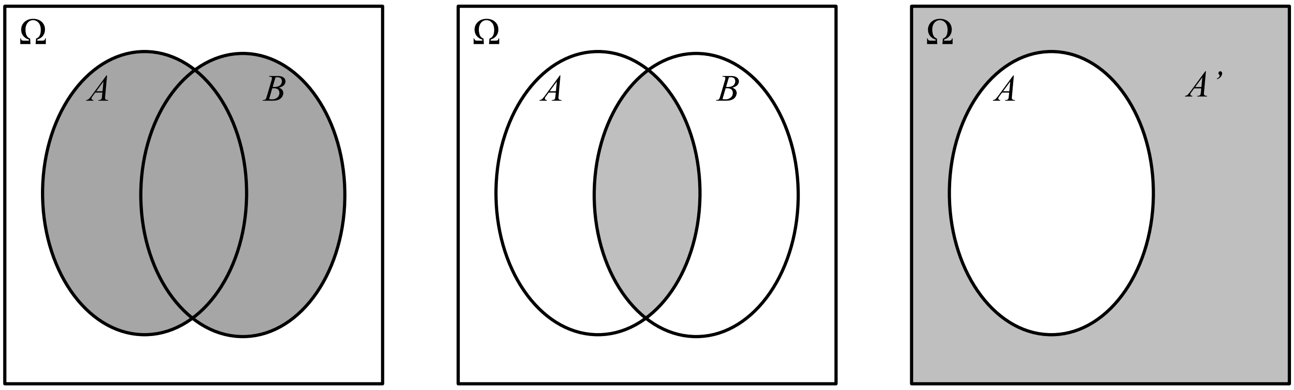
4. *Δύο σύνολα είναι ίσα όταν περιέχουν ακριβώς τα ίδια στοιχεία (ή, ισοδύναμα, το ένα είναι υποσύνολο του άλλου και αντίστροφα.)*

5. *Το* A *είναι γνήσιο υποσύνολο του* B *αν το* $A \subseteq B$, *αλλά τα* A, B *δεν είναι ίσα, δηλαδή υπάρχει στοιχείο του* B *που δεν ανήκει στο* A. *Γράφουμε* $A \subset B$.

6. *Το κενό σύνολο* $\emptyset$ *ή* $\{\}$ *έχει την ιδιότητα ότι δεν περιέχει κανένα στοιχείο, δηλαδή* $\alpha \notin \emptyset$ *για οποιοδήποτε* α.

7. *Η ένωση* $A \cup B$ *δύο συνόλων* A, B *είναι το σύνολο που αποτελείται από όλα τα στοιχεία που ανήκουν στο* A *ή στο* B *(ή και στα δύο). Δείτε το Σχήμα 1.1. Γενικότερα, η ένωση ενός πεπερασμένου πλήθους συνόλων* $A_1, A_2, \dots, A_N$ *συμβολίζεται ως*

$$A_1 \cup A_2 \cup \dots \cup A_N = \bigcup_{i=1}^{N} A_i$$



Σχήμα 1.1: Γραφική αναπαράσταση της ένωσης $A \cup B$, της τομής $A \cap B$, και του συμπληρώματος A'. Τέτοια σχήματα καλούνται Διαγράμματα Venn και είναι πολύ χρήσιμα στην διαισθητική ερμηνεία πολλών ιδιοτήτων των συνόλων.

και περιέχει όλα τα στοιχεία του A_1, όλα τα στοιχεία του A_2, κ.ο.κ., και μόνο αυτά. Συνεπώς, ένα στοιχείο x ανήκει στην ένωση αν και μόνο αν ανήκει σε τουλάχιστον ένα από τα A_i. Η ένωση $\bigcup_{i=1}^{\infty} A_i$ μιας ακολουθίας άπειρου πλήθους συνόλων $A_1, A_2, \ldots$ ορίζεται ανάλογα.

8. *Η τομή $A \cap B$ δύο συνόλων A, B είναι το σύνολο που αποτελείται από όλα τα στοιχεία που ανήκουν και στο A και στο B. Δείτε το Σχήμα 1.1. Γενικότερα, η τομή ενός πεπερασμένου πλήθους συνόλων $A_1, A_2, \ldots, A_N$ συμβολίζεται ως*

$$A_1 \cap A_2 \cap \cdots \cap A_N = \bigcap_{i=1}^{N} A_i$$

και αποτελείται ακριβώς από αυτά τα στοιχεία που περιέχονται σε όλα τα A_i. Η τομή $\bigcap_{i=1}^{\infty} A_i$ μιας ακολουθίας άπειρου πλήθους συνόλων $A_1, A_2, \ldots$ ορίζεται ανάλογα.

9. *Έστω κάποιο βασικό ή καθολικό σύνολο Ω, στο οποίο υποθέτουμε ότι βρίσκονται όλα τα στοιχεία που μπορούν να υπάρξουν στα πλαίσια κάποιου μαθηματικού μοντέλου. Το συμπλήρωμα A' ενός συνόλου A που είναι υποσύνολο του βασικού συνόλου Ω αποτελείται από όλα τα στοιχεία του Ω που δεν ανήκουν στο A. Δείτε το Σχήμα 1.1.*

Παρατήρηση: Εναλλακτικά, για το συμπλήρωμα ενός συνόλου A χρησιμοποιούνται και οι συμβολισμοί $\bar{A}$ και A^c (από το complement).

Λήμμα 1.1. (Βασικές ιδιότητες συνόλων)
Έστω οποιαδήποτε σύνολα A, B, C. Έχουμε:

1. $A \cup B = B \cup A$ *(η ένωση είναι αντιμεταθετική).*
2. $A \cap B = B \cap A$ *(η τομή είναι αντιμεταθετική).*
3. $(A')' = A$.
4. $A \cup A' = \Omega$, *όπου* Ω *είναι το καθολικό σύνολο.*
5. $A \cap A' = \emptyset$.
6. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *(η τομή είναι επιμεριστική ως προς την ένωση).*
7. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *(η ένωση είναι επιμεριστική ως προς την τομή).*
8. $(A \cup B)' = A' \cap B'$ *(Νόμος De Morgan).*
9. $(A \cap B)' = A' \cup B'$ *(Νόμος De Morgan).*

Απόδειξη. Θα δείξουμε την απόδειξη μόνο της τελευταίας ιδιότητας, $(A \cap B)' = A' \cup B'$, καθώς οι αποδείξεις των πρώτων 5 είναι προφανείς, και οι αποδείξεις των 6,7,8 ανάλογες με την απόδειξη της τελευταίας. Η μέθοδος που θα ακολουθήσουμε για να αποδείξουμε ότι τα δύο σύνολα είναι ίσα είναι η κλασική: θα αποδείξουμε ότι το ένα είναι υποσύνολο του άλλου και το αντίστροφο.

Έστω λοιπόν $x \in (A \cap B)'$. Το x δεν ανήκει στο $A \cap B$, άρα είτε δεν ανήκει στο A, είτε δεν ανήκει στο B, αλλιώς έχουμε άτοπο. Άρα, είτε ανήκει στο A', είτε ανήκει στο B'. Άρα ανήκει στην ένωση $A' \cup B'$. Άρα αποδείξαμε πως $(A \cap B)' \subseteq A' \cup B'$.

Έστω τώρα $x \in A' \cup B'$. Τότε παίρνουμε δύο περιπτώσεις:

1. Αν το x ανήκει στο A, τότε ανήκει στο B', αλλιώς έχουμε άτοπο. Άρα δεν ανήκει στο B, άρα ούτε και στο $A \cap B$, άρα ανήκει στο $(A \cap B)'$.

2. Αν πάλι το x δεν ανήκει στο A, τότε δεν ανήκει και στο $A \cap B$, άρα ανήκει στο $(A \cap B)'$.

Οπότε, σε κάθε περίπτωση, το x ανήκει στο $(A \cap B)'$, και αποδείξαμε πως $A' \cup B' \subseteq (A \cap B)'$.

Αφού λοιπόν έχουμε ότι $(A \cap B)' \subseteq A' \cup B'$ και $A' \cup B \subseteq (A \cap B)'$, αναγκαστικά $(A \cap B)' = A' \cup B'$.

Μπορείτε να γράψετε τις αποδείξεις των ιδιοτήτων 6,7,8; Οι περισσότερες από τις άνω ιδιότητες γενικεύονται και στην περίπτωση περισσότερων των 2 ή και απείρων συνόλων. Μπορείτε να κάνετε μερικές γενικεύσεις; Επίσης, μπορείτε να σκεφτείτε και άλλες ιδιότητες; □

Παράδειγμα 1.1. Παρατηρήστε πως

$$\bigcup_{n=1}^{\infty}[0,n]=[0,\infty).$$

Πράγματι, έστω ένας οποιοσδήποτε αριθμός $x \in [0,\infty)$. Έστω κάποιο $n_0 > x$. Τότε ο $x \in [0,n_0]$, άρα $x \in \bigcup_{n=1}^{\infty}[0,n]$. Αντιστρόφως, αν $x \in \bigcup_{n=1}^{\infty}[0,n]$ τότε θα ανήκει σε κάποιο $[0,n_0]$, άρα και στο $[0,\infty)$. Άρα πράγματι τα δύο σύνολα ταυτίζονται.

Παράδειγμα 1.2. *(Άπειρη ένωση και άπειρη τομή)* Έστω σύνολα $A_n = \big[0, 1-\frac{1}{n}\big)$, $n = 1,2,\ldots$, και $B_n = \big[0, 1+\frac{1}{n}\big)$, $n = 1,2,\ldots$ Έστω τα σύνολα

$$\bigcup_{n=1}^{\infty} A_n, \quad \bigcap_{n=1}^{\infty} A_n, \quad \bigcup_{n=1}^{\infty} B_n, \quad \bigcap_{n=1}^{\infty} B_n.$$

Μπορείτε να περιγράψετε αυτά τα σύνολα πιο απλά; Μπορείτε να αποδείξετε αυστηρά τον ισχυρισμό σας;

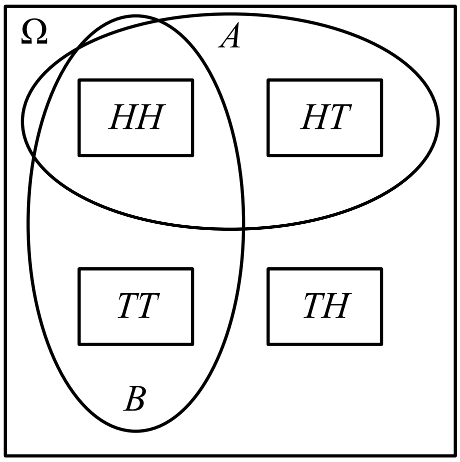
1.2 Δειγματικοί Χώροι

Ορισμός 1.2. (Δειγματικοί χώροι)

1. *Ο δειγματικός χώρος (Δ.Χ.) Ω είναι το σύνολο όλων των δυνατών αποτελεσμάτων ενός τυχαίου πειράματος.*
2. *Τα στοιχεία του δειγματικού χώρου καλούνται αποτελέσματα.*
3. *Οποιοδήποτε υποσύνολο $A \subseteq \Omega$ του δειγματικού χώρου Ω ονομάζεται ενδεχόμενο.*
4. *Τα ενδεχόμενα που αποτελούνται από ένα μόνο αποτέλεσμα, δηλαδή τα υποσύνολα $A \subseteq \Omega$ της μορφής $A = \{\omega\}$ όπου $\omega \in \Omega$, λέγονται στοιχειώδη ενδεχόμενα.*
5. *Δύο ενδεχόμενα A, B είναι ξένα (μεταξύ τους) όταν δεν έχουν κανένα κοινό στοιχείο, δηλαδή $A \cap B = \emptyset$.*

Παρατηρήσεις

1. Μερικές φορές για τον δειγματικό χώρο χρησιμοποιείται το σύμβολο S (από το sample space).
2. Παρατηρούμε ότι ο δειγματικός χώρος μπορεί πάντα να εκφραστεί ως η ένωση τόσων αποτελεσμάτων όσα τα αποτελέσματα που περιέχει. Π.χ., αν $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$, τότε $\Omega = \{\omega_1\} \cup \{\omega_2\} \cup \cdots \cup \{\omega_N\}$. Γενικότερα, κάθε ενδεχόμενο μπορεί να εκφραστεί ως η ένωση τόσων στοιχειωδών ενδεχομένων όσα τα αποτελέσματα που περιέχει.
3. Επίσης παρατηρούμε ότι δύο οποιαδήποτε αποτελέσματα $\{\omega_1\}$ και $\{\omega_2\}$ είναι ξένα μεταξύ τους (αρκεί, βεβαίως, να μην είναι τα ίδια, δηλαδή το στοιχείο ω_1 να είναι διαφορετικό απ' το ω_2).
4. Διαισθητικά:
   (α′) Ένα ενδεχόμενο A είναι ένα σύνολο αποτελεσμάτων για το οποίο μας ενδιαφέρει αν θα συμβεί ένα από τα αποτελέσματα που το απαρτίζουν ή όχι, χωρίς να μας ενδιαφέρει ποιο από όλα τα αποτελέσματά του θα συμβεί.
   (β′) Το συμπλήρωμα του A' περιγράφει το αντίθετο ενδεχόμενο, δηλαδή την περίπτωση του να *μην* συμβεί κανένα από τα αποτελέσματα του A.
   (γ′) Παρομοίως, η ένωση $A \cup B$ δύο ενδεχομένων A, B είναι το ενδεχόμενο του να συμβεί *το A ή το B*, και η τομή τους $A \cap B$ είναι το ενδεχόμενο να συμβούν *και τα δύο*.



Σχήμα 1.2: Γραφική αναπαράσταση των διάφορων ενδεχόμενων του Παραδείγματος 1.4.

(δ′) Διαισθητικά, τα ενδεχόμενα A, B είναι ξένα αν είναι αδύνατον να συμβούν συγχρόνως.

5. Αν και τυπικά οι έννοιες του αποτελέσματος και του στοιχειώδους ενδεχόμενου δεν ταυτίζονται (γιατί;), εντούτοις συχνά θα τις αντιμετωπίζουμε ως ισοδύναμες.

Παράδειγμα 1.3. *(Ρίψη κέρματος)* Κατά μια έννοια το πιο απλό πείραμα που μπορεί να υπάρξει είναι η ρίψη ενός κέρματος. Ο δειγματικός χώρος είναι ο

$$\Omega = \{H, T\},$$

όπου το H σημαίνει κορώνα (heads), και το T σημαίνει γράμματα (tails). Τα δυνατά διακριτά ενδεχόμενα είναι μόλις 4:

$$\emptyset,\ \{H\},\ \{T\},\ \Omega.$$

Παράδειγμα 1.4. *(Ρίψη δύο κερμάτων)* Ρίχνουμε ένα κέρμα 2 φορές. Εδώ το σύνολο όλων των δυνατών αποτελεσμάτων είναι το

$$\Omega = \{HH, HT, TH, TT\}.$$

Μπορείτε να υπολογίσετε πόσα είναι τα διακριτά ενδεχόμενα; Έστω A το ενδεχόμενο του να έρθει Κ την πρώτη φορά, και B το ενδεχόμενο να έρθει το ίδιο αποτέλεσμα δύο φορές, δηλαδή,

$$A = \{HH, HT\}, \quad B = \{HH, TT\}.$$

Ποια είναι η ένωσή τους; Πως θα την περιγράφατε με λόγια; Ποια η τομή τους; Στο Σχήμα 1.2 απεικονίζουμε τον Ω και τα A, B.

Παράδειγμα 1.5. *(Τυχαία παιδιά)* Έστω πως εκτελείται το ακόλουθο πείραμα: ένα ζευγάρι κάνει n παιδιά, καθένα εκ των οποίων μπορεί να είναι αγόρι ή κορίτσι. Ο δειγματικός χώρος είναι το σύνολο Ω που αποτελείται από όλες τις δυνατές n-άδες της μορφής $XX\ldots X$, όπου $X = B$ (boy) ή $X = G$ (girl). Άρα περιέχει 2^n τέτοια αποτελέσματα. Π.χ. το αποτέλεσμα $\omega = GBB\ldots B$ περιγράφει την περίπτωση όπου το ζευγάρι έκανε πρώτα ένα κορίτσι και μετά $(n-1)$ αγόρια.

Από ποια (και πόσα) αποτελέσματα αποτελείται το ενδεχόμενο να έχει η οικογένεια το πολύ 1 αγόρια; Από ποια (και πόσα) αποτελέσματα αποτελείται το ενδεχόμενο να έχει η οικογένεια το πολύ 2 αγόρια, αν $n \geq 2$;

Παράδειγμα 1.6. *(Κι άλλα τυχαία παιδιά)* Έστω πως εκτελείται το ακόλουθο πείραμα: ένα ζευγάρι κάνει παιδιά επ' άπειρο, μέχρι να κάνει το πρώτο κορίτσι, και μετά σταματάει. Ο δειγματικός χώρος Ω εδώ περιέχει δύο ειδών στοιχεία: εκείνα που αντιστοιχούν στα αποτελέσματα πεπερασμένου μήκους G, BG, BBG, $BBBG$, κ.ο.κ., και ένα που αντιστοιχεί στην περίπτωση όπου το ζευγάρι κάνει άπειρα αγόρια, το $BBBBBB\ldots$ (Το αποτέλεσμα BG αντιστοιχεί στο να κάνει το ζευγάρι πρώτα ένα αγόρι (boy) και μετά ένα κορίτσι (girl)). Δηλαδή,

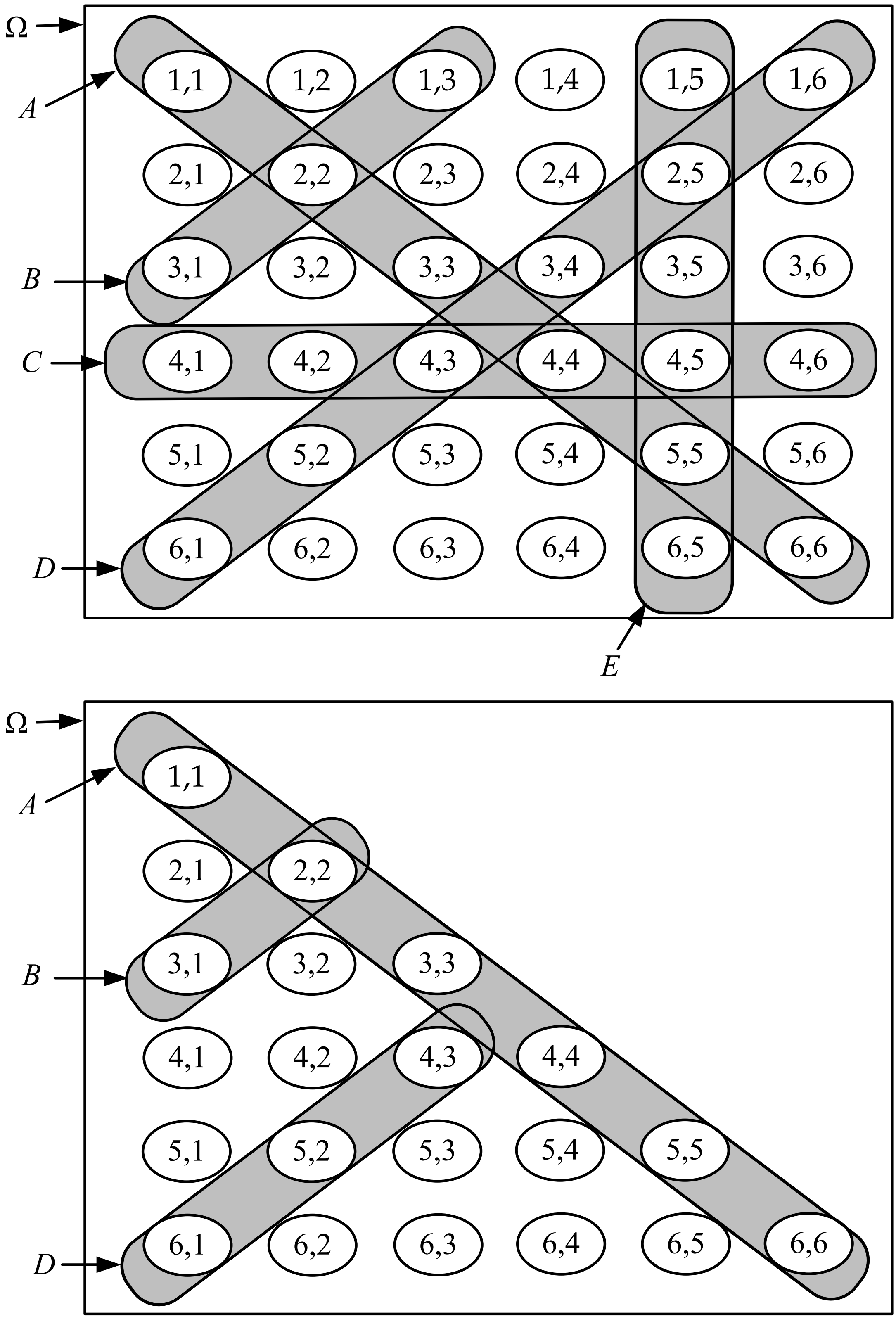
$$\Omega = \{G,\ \ BG,\ \ BBG,\ \ \ldots,\ \ B\ldots BG,\ \ \ldots,\ \ BBBBBB\ldots\}.$$

Από ποια αποτελέσματα αποτελείται το ενδεχόμενο να έχει η οικογένεια το πολύ 3 παιδιά;

Παράδειγμα 1.7. *(Δύο διαδοχικές ζαριές)* Ρίχνουμε ένα ζάρι 2 φορές και καταγράφουμε τα δύο αποτελέσματα με τη σειρά που ήρθαν. Ο δειγματικός χώρος έχει σχεδιαστεί στο Σχήμα 1.3 (άνω). Επίσης, στο ίδιο σχήμα έχουμε σχεδιάσει τα

1. $A = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$ =«Φέραμε διπλές» .
2. $B = \{(1,3),(2,2),(3,1)\}$ =«Άθροισμα 4».
3. $C = \{(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)\}$ =«Πρώτο ζάρι 4».
4. $D = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$ =«Άθροισμα 7».
5. $E = \{(1,5),(2,5),(3,5),(4,5),(5,5),(6,5)\}$ =«Δεύτερο ζάρι 5».

Παράδειγμα 1.8. *(Δύο ταυτόχρονες ζαριές)* Έστω το Παράδειγμα 1.7 όπου τα ζάρια ρίχνονται ταυτόχρονα, και δεν είμαστε σε θέση να τα ξεχωρίζουμε μεταξύ τους. Σε αυτή την περίπτωση, οι ζαριές που προκύπτουν δεν είναι διατεταγμένες δυάδες, και συνεπώς δεν έχει νόημα να υπάρχουν στον δειγματικό χώρο ταυτόχρονα οι δυάδες $(1,2)$ και $(2,1)$, $(1,3)$ και $(3,1)$, κ.ο.κ. Ο νέος δειγματικός χώρος Ω και τα τροποποιημένα ενδεχόμενα A, B, D εμφανίζονται στο Σχήμα 1.3 (κάτω). Τα ενδεχόμενα C, E άνω δεν έχουν νόημα σε αυτό το δειγματικό χώρο.



Σχήμα 1.3: Παραδείγματα 1.7 και 1.8.

Παρατηρήσεις

1. Όπως φαίνεται από τα άνω παραδείγματα, γενικώς μπορούμε να περιγράψουμε τα ενδεχόμενα με δύο τρόπους: είτε με απαρίθμηση των αποτελεσμάτων τους, είτε περιφραστικά.

2. Όπως φαίνεται από τα τελευταία δύο παραδείγματα, σε πολλές περιπτώσεις πειραμάτων έχουμε πολλές διαφορετικές επιλογές για το ποιος θα είναι ο δειγματικός χώρος. Η επιλογή του δειγματικού χώρου είναι σε μεγάλο βαθμό θέμα μοντελοποίησης, και εξαρτάται από το τι θέλουμε να υπολογίσουμε.

Παράδειγμα 1.9. *(Δειγματικός χώρος με άπειρα αποτελέσματα)* Έστω ζυγαριά που μπορεί να μετρήσει βάρη μέχρι 100 κιλά. Αν το τυχαίο πείραμα είναι η ζύγιση ενός ατόμου, ο δειγματικός χώρος είναι ο $\Omega = [0, 100]$, και προφανώς αποτελείται από άπειρα, και μάλιστα μη αριθμήσιμα αποτελέσματα. Το ενδεχόμενο το βάρος ενός ατόμου να είναι το πολύ 50 κιλά είναι το $[0, 50]$.

Παράδειγμα 1.10. *(Πραγματικό γεγονός)* Έστω πως ρίχνουμε ένα βελάκι σε ένα στόχο με σχήμα κύκλου, και διάμετρο 40 cm. Αν πετύχουμε το στόχο το βελάκι μένει καρφωμένο, και αν αστοχήσουμε το βελάκι πέφτει στο πάτωμα και το κλέβει ο σκύλος μας. Εδώ ο δειγματικός χώρος Ω μπορεί να περιγραφεί ως το ακόλουθο σύνολο:

$$\Omega = \Big\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 40^2\Big\} \cup \{\text{ΣΚΥΛΟΣ}\}.$$

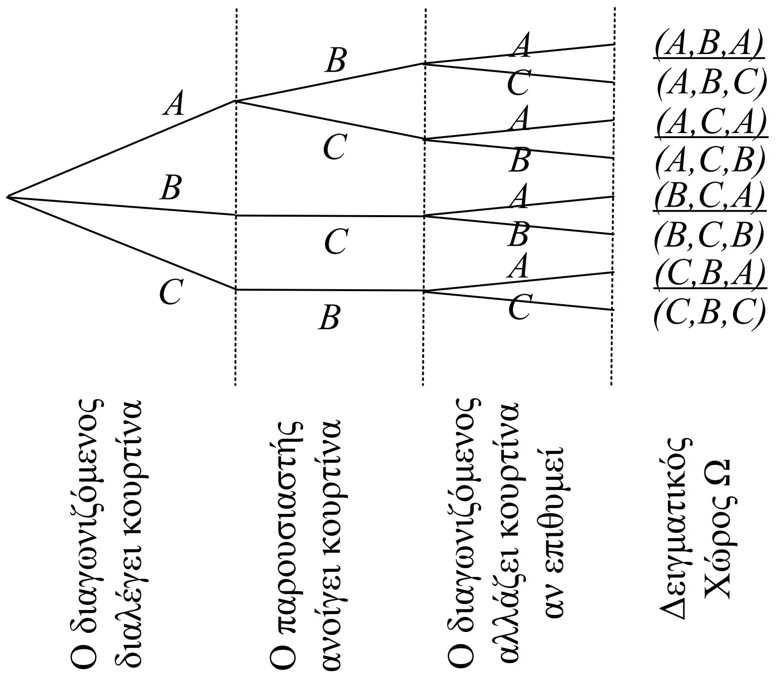
Παρατηρήσεις

1. Όπως φαίνεται από το προηγούμενο παράδειγμα, μερικοί δειγματικοί χώροι μπορεί να έχουν πολύ περίεργη δομή.

2. Σε πολλές περιπτώσεις, ο δειγματικός χώρος μπορεί να πάρει τη μορφή δέντρου. Δείτε το επόμενο παράδειγμα.

Παράδειγμα 1.11. *(Monty Hall)* Έστω το ακόλουθο τηλεπαιχνίδι: ο διαγωνιζόμενος επιλέγει μια από τρεις κουρτίνες, αφού του πουν πως μία από αυτές κρύβει ένα δώρο και οι άλλες δύο δεν κρύβουν τίποτα (χωρίς, φυσικά, να του πουν που είναι το δώρο). Αφού διαλέξει, ο παρουσιαστής του δείχνει μια απ' τις άλλες δύο κουρτίνες που είναι κενή, και δίνει την επιλογή στον διαγωνιζόμενο να κρατήσει την αρχική του επιλογή ή να επιλέξει την κουρτίνα που παραμένει κρυφή. Ο διαγωνιζόμενος επιλέγει, και το παιχνίδι τελειώνει, είτε με νίκη του διαγωνιζόμενου (αν η τελική κουρτίνα που επέλεξε περιέχει το δώρο), είτε με ήττα του διαγωνιζόμενου (αν η κουρτίνα που επέλεξε δεν περιέχει το δώρο).[1]

[1] Αυτό το παιχνίδι ήταν επί χρόνια τηλεπαιχνίδι στην Αμερική, γνωστό με το όνομα «Monty Hall». Το βασικό ερώτημα, το οποίο θα εξετάσουμε αργότερα, είναι ποια είναι η πιο συμφέρουσα στρατηγική για τον παίκτη – να κρατήσει την αρχική του κουρτίνα, ή να αλλάξει;



Σχήμα 1.4: Ο δειγματικός χώρος του Παραδείγματος 1.11. Τα αποτελέσματα που ανήκουν στο ενδεχόμενο της νίκης είναι υπογραμμισμένα.

Πως μπορούμε να περιγράψουμε τον δειγματικό χώρο; Μια επιλογή είναι η ακόλουθη. Έστω πως ονομάζουμε κουρτίνα A την κουρτίνα όπου βρίσκεται το δώρο, και κουρτίνες B, C τις άλλες δύο. Μπορούμε να περιγράψουμε τα αποτελέσματα ως τριάδες της μορφής (X, X, X), όπου τα X παίρνουν τιμές A, B, ή C, και το πρώτο στοιχείο δείχνει την επιλογή του διαγωνιζόμενου, το δεύτερο την κουρτίνα που αποκαλύφθηκε, και το τρίτο την κουρτίνα που επέλεξε τελικά ο διαγωνιζόμενος.

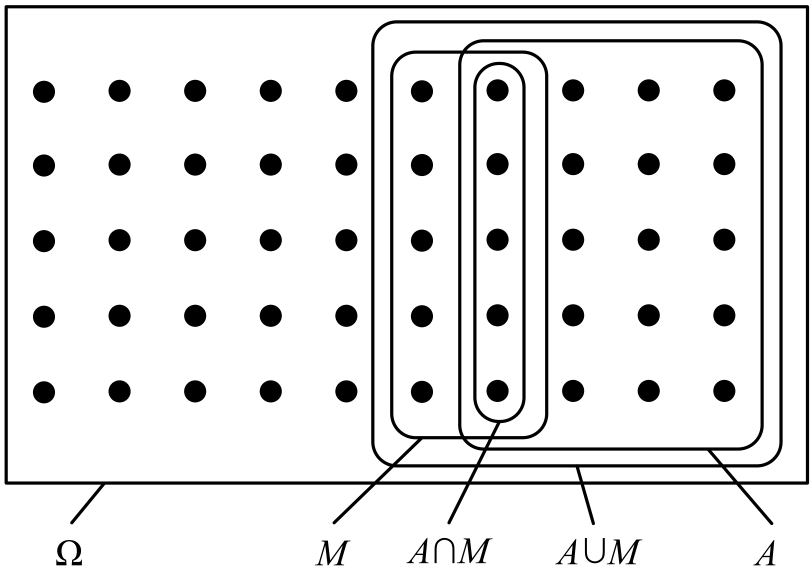
Προφανώς υπάρχουν 3 επιλογές για το πρώτο στοιχείο. Αλλά για το δεύτερο στοιχείο υπάρχουν 2 επιλογές αν ο διαγωνιζόμενος έχει αρχικά επιλέξει την κουρτίνα με το δώρο, ενώ υπάρχει μόνο μία αν ο διαγωνιζόμενος έχει επιλέξει κενή κουρτίνα. Για το τρίτο στοιχείο, υπάρχουν πάντα δύο επιλογές. Ο αντίστοιχος δειγματικός χώρος έχει σχεδιαστεί στο Σχήμα 1.4.

Αν θέλουμε τώρα να ορίσουμε, π.χ., το ενδεχόμενο $N =$«ο παίκτης κέρδισε το δώρο», παρατηρούμε πως τα αποτελέσματα που καταλήγουν σε νίκη για τον διαγωνιζόμενο είναι εκείνα που έχουν τελευταίο στοιχείο το A, δηλαδή, $N = \{ABA, ACA, BCA, CBA\}$.

Παρατηρήσεις

1. Σε ποια άλλα από τα προηγούμενα παραδείγματα ο δειγματικός χώρος έχει επίσης τη μορφή δέντρου;

2. Σε πολλά προβλήματα, μας δίνονται τα πλήθη των στοιχείων κάποιων ενδεχόμενων,



Σχήμα 1.5: Γραφική αναπαράσταση των ενδεχόμενων στο Παράδειγμα 1.12.

και καλούμαστε να βρούμε τα πλήθη των στοιχείων κάποιων άλλων. Δείτε για παράδειγμα το ακόλουθο παράδειγμα.

3. Με $|A|$ θα συμβολίσουμε το πλήθος των στοιχείων ενός συνόλου A.

Παράδειγμα 1.12. *(Φοιτητές)* Από 50 φοιτητές που βρίσκονται σε μια αίθουσα, οι 20 έχουν αυτοκίνητο, οι 10 έχουν μοτοσυκλέτα, και οι 25 δεν έχουν κανένα απ' τα δύο. Επιλέγουμε έναν φοιτητή στην τύχη.

Εδώ μπορούμε να ορίσουμε τα εξής ενδεχόμενα, που εμφανίζονται στο Σχήμα 1.5.

$$\begin{aligned} \Omega &= \text{Όλοι οι φοιτητές}, \\ A &= \text{Όσοι έχουν αυτοκίνητο}, \\ M &= \text{Όσοι έχουν μοτοσυκλέτα}, \\ A \cup M &= \text{Όσοι έχουν τουλάχιστον το ένα απ' τα δύο μέσα}, \\ A \cap M &= \text{Όσοι έχουν και τα δύο}. \end{aligned}$$

Έχει δοθεί ότι $|\Omega| = 50$, $|A| = 20$, $|M| = 10$, και $|(A \cup M)'| = 25$. Θα βρούμε πόσοι φοιτητές είναι στο $A \cup M$ και πόσοι φοιτητές είναι στο $A \cap M$.

Για το $A \cup M$, εύκολα υπολογίζουμε ότι

$$|A \cup M| = |\Omega| - |(A \cup M)'| = 50 - 25 = 25.$$

Για το $A \cap M$, από την γραφική αναπαράσταση στο Σχήμα 1.5, παρατηρούμε πως $|A \cup M| = |A| + |M| - |A \cap M|$, όπου αφαιρούμε τα στοιχεία του συνόλου $A \cap M$ για να μην μετρηθούν δύο φορές. Παρατηρούμε επίσης ότι έχουμε βρει $|A \cup M| = 25$, ενώ μας δίνεται και ότι $|A| = 20$ και $|M| = 10$, άρα, $|A \cap M| = 20 + 10 - 25 = 5$.

1.3 Μέτρο Πιθανότητας

Είμαστε έτοιμοι να δώσουμε έναν αυστηρά μαθηματικό ορισμό της έννοιας της πιθανότητας. Αν και, εκ πρώτης όψεως, ο ορισμός φαίνεται δυσνόητος και πολύ απομακρυσμένος απ΄ αυτό που διαισθητικά ονομάζουμε «πιθανότητα», όπως θα δούμε στα παραδείγματα που ακολουθούν, στην πράξη είναι πολύ απλός και εύχρηστος.

Ορισμός 1.3. (Μέτρο πιθανότητας) *Έστω ένας δειγματικός χώρος Ω και έστω $\mathcal{F}$ το δυναμοσύνολο του Ω, δηλαδή το σύνολο που έχει ως στοιχεία όλα τα ενδεχόμενα $A \subset \Omega$ (συμπεριλαμβανομένου και του κενού συνόλου $\emptyset$). Ένα μέτρο πιθανότητας είναι μια συνάρτηση $P : \mathcal{F} \to [0, 1]$ η οποία ικανοποιεί τις παρακάτω ιδιότητες, που καλούνται αξιώματα των πιθανοτήτων:*

1. *(Πρώτο Αξίωμα) $P(A) \geq 0$ για οποιοδήποτε ενδεχόμενο $A \in \mathcal{F}$.*
2. *(Δεύτερο Αξίωμα) $P(\Omega) = 1$.*
3. *(Τρίτο Αξίωμα) Αν δύο ενδεχόμενα $A, B \in \mathcal{F}$ είναι ξένα (δηλαδή $A \cap B = \emptyset$), τότε*
$$P(A \cup B) = P(A) + P(B).$$
Και γενικότερα, αν $A_1, A_2, \ldots$ είναι μια οποιαδήποτε (πεπερασμένη ή όχι) ακολουθία ξένων ενδεχομένων (δηλαδή $A_i \cap A_j = \emptyset$ για κάθε $i \neq j$), τότε
$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots.$$

Παρατηρήσεις

1. Τα πρώτα δύο αξιώματα είναι ιδιότητες που προφανώς πρέπει να έχουν οι πιθανότητες. Πράγματι, δεν μπορεί η πιθανότητα να συμβεί κάτι να είναι αρνητική. Επιπλέον, η πιθανότητα να συμβεί *οτιδήποτε* πρέπει να είναι 100%, δηλαδή 1.

2. Το τρίτο αξίωμα είναι και αυτό λογικό, μετά από λίγη σκέψη. Για παράδειγμα, έστω πως ρίχνουμε ένα ζάρι, και έστω ο δειγματικός χώρος $\Omega = \{1, 2, 3, 4, 5, 6\}$. Έστω πως
$$\begin{aligned} P(\{1,2\}) &= 1/3, \\ P(\{3,4\}) &= 1/3, \\ P(\{2,3\}) &= 1/3. \end{aligned}$$
Τότε, *πρέπει* να ισχύει
$$P(\{1,2,3,4\}) = P(\{1,2\} \cup \{3,4\}) = P(\{1,2\}) + P(\{3,4\}) = 1/3 + 1/3 = 2/3,$$

αλλά *δεν είναι απαραίτητο* να ισχύει

$$P(\{1,2,3\}) = P(\{1,2\} \cup \{2,3\}) = P(\{1,2\}) + P(\{2,3\}) = 1/3 + 1/3 = 2/3.$$

Μπορείτε να βρείτε ένα άνω και ένα κάτω φράγμα για την άνω πιθανότητα;

3. Ο άνω ορισμός είναι αρκετός για τις ανάγκες του μαθήματος. Πρέπει πάντως να αναφερθεί ότι σε ορισμένες περιπτώσεις δεν είναι σωστός, καθώς μπορεί να αποδειχτεί ότι αν το Ω είναι μη αριθμήσιμο, τότε δεν υπάρχει μέτρο πιθανότητας που να μπορεί να οριστεί στο δυναμοσύνολο $\mathcal{F}$, και αναγκαστικά πρέπει να περιοριστούμε σε ένα σύνολο συνόλων με μικρότερο μέγεθος. Το συγκεκριμένο πρόβλημα είναι αυστηρώς μεταπτυχιακού επιπέδου, και μπορείτε να το αγνοήσετε.

4. Μια μεγάλη κατηγορία προβλημάτων μας δίνει τις πιθανότητες κάποιων ενδεχόμενων, και ζητά να υπολογίσουμε τις πιθανότητες άλλων. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 1.13. *(Τρία δίκτυα)* Σε ένα εργαστήριο πληροφορικής λειτουργούν τρία δίκτυα, από τα οποία έχουμε παρατηρήσει πως στο πρόσφατο παρελθόν

- 30% των ημερών, τουλάχιστον ένα δίκτυο *δεν* λειτουργεί,
- 10% των ημερών, ακριβώς δύο δεν λειτουργούν,
- 5% των ημερών, δεν λειτουργεί κανένα δίκτυο.

Έστω το τυχαίο πείραμα της παρατήρησης του ποια δίκτυα λειτουργούν σήμερα. Επιλέγουμε τον δειγματικό χώρο

$$\Omega = \{000, 001, 010, 011, 100, 101, 110, 111\},$$

όπου, για παράδειγμα, το αποτέλεσμα 011 σημαίνει ότι λειτουργούν το δεύτερο και το τρίτο δίκτυο.

Έστω B_i το ενδεχόμενο του να λειτουργούν ακριβώς i απ΄ τα τρία δίκτυα, για $i = 0, 1, 2, 3$. Για παράδειγμα, το $B_2 = \{110, 101, 011\}$ είναι το ενδεχόμενο να λειτουργούν ακριβώς δύο δίκτυα. Παρατηρήστε πως $\Omega = \bigcup_{i=0}^{3} B_i$.

Από τις υποθέσεις μας δικαιολογούμαστε να θέσουμε ότι $P(B_0) = 5\% = 0.05$ και $P(B_1) = 10\% = 0.1$. Επιπλέον, η πρώτη υπόθεση μας λέει ότι $P(B_3') = 30\% = 0.3$ (γιατί;). Αλλά ποιες είναι οι πιθανότητες των B_2 και B_3;

Για το B_3 παρατηρούμε ότι, εφόσον το «λειτουργούν και τα τρία δίκτυα» είναι το αντίθετο του «τουλάχιστον ένα δεν λειτουργεί», διαισθητικά περιμένουμε να ισχύει ότι

$$P(B_3) = 1 - P(B_3') = 1 - 0.3 = 0.7.$$

Πράγματι, μπορούμε να δούμε και αυστηρά πως αφού τα B_3 και B_3' είναι ξένα και επιπλέον $B_3 \cup B_3' = \Omega$, από το τρίτο αξίωμα θα πρέπει να έχουμε

$$P(\Omega) = P(B_3 \cup B_3') = P(B_3) + P(B_3') \Rightarrow P(\Omega) = P(B_3) + P(B_3')$$
$$\Rightarrow 1 = P(B_3) + P(B_3') \Rightarrow P(B_3) = 1 - P(B_3').$$

Η δεύτερη συνεπαγωγή προέκυψε από το δεύτερο αξίωμα.

Για το B_2 τώρα, παρατηρούμε ότι το ενδεχόμενο B_3' του να μην λειτουργεί τουλάχιστον ένα δίκτυο μπορεί να εκφραστεί ως η ένωση

$$B_3' = B_0 \cup B_1 \cup B_2$$

όπου τα B_0, B_1 και B_2 είναι εξ ορισμού ξένα. Άρα,

$$0.3 = P(B_3') = P(B_0) + P(B_1) + P(B_2) = 0.05 + 0.1 + P(B_2),$$

οπότε βρίσκουμε πως $P(B_2) = 0.15$ ή 15%.

Παρατηρήσεις

1. Παρατηρήστε ότι χρησιμοποιήσαμε την πολύ κοινή σύμβαση ότι όταν η πιθανότητα κάποιου ενδεχόμενου $P(A) = a$ γράφουμε και $P(A) = 100a\%$.

2. Στο άνω παράδειγμα, όπως και σε πολλά άλλα που θα δούμε στη συνέχεια, υπήρχαν και άλλοι τρόποι για να ορίσουμε τον δειγματικό χώρο. Για παράδειγμα, θα μπορούσαμε να θέσουμε
$$\Omega = \{0, 1, 2, 3\},$$
όπου ο αριθμός δείχνει απλώς τα δίκτυα που λειτουργούν. Πως θα περιγράφαμε τότε τα ενδεχόμενα B_i;

3. Επιπλέον, θα μπορούσαμε να υπολογίσουμε τις ζητούμενες πιθανότητες ορίζοντας μεν τα ενδεχόμενα B_i με λόγια, αλλά χωρίς να ορίσουμε ρητώς τον δειγματικό χώρο. Στη συνέχεια, πολύ συχνά θα αποφεύγουμε να ορίζουμε ρητώς τον δειγματικό χώρο, όταν αυτό δεν θα είναι απαραίτητο για την επίλυση του παραδείγματος.

4. Αργότερα θα αποδείξουμε αρκετές ιδιότητες των πιθανοτήτων που κάνουν υπολογισμούς όπως αυτούς του προηγούμενου παραδείγματος αρκετά γρηγορότερους.

Παράδειγμα 1.14. *(Δίκαιο ζάρι)* Ρίχνουμε ένα «δίκαιο» ζάρι. Ο δειγματικός χώρος $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Εφόσον το ζάρι είναι «δίκαιο», απαιτούμε το αντίστοιχο μέτρο πιθανότητας P που περιγράφει αυτό το πείραμα να δίνει την ίδια πιθανότητα, δηλαδή $1/6$, σε κάθε δυνατό στοιχειώδες ενδεχόμενο, δηλαδή να ισχύει $P(\{i\}) = 1/6$ για κάθε $i = 1, \ldots, 6$. Οι

πιθανότητες των υπολοίπων ενδεχόμενων μπορούν να προκύψουν με χρήση του τρίτου αξιώματος, αφού όλα τους μπορούν να γραφούν σαν ένωση ξένων στοιχειωδών ενδεχόμενων.

Για παράδειγμα, ορίζουμε τα ενδεχόμενα

$$\begin{aligned} A &= \text{«ζυγό αποτέλεσμα»} = \{2,4,6\}, \\ B &= \text{«μονό αποτέλεσμα»} = \{1,3,5\}. \end{aligned}$$

Για να υπολογίσουμε την πιθανότητα του ενδεχομένου A (η οποία διαισθητικά είναι «προφανώς» ίση με $1/2$), παρατηρούμε πως

$$A = \{2\} \cup \{4\} \cup \{6\}$$

και πως όλα τα στοιχειώδη ενδεχόμενα είναι ξένα μεταξύ τους. Άρα από το τρίτο αξίωμα έχουμε

$$P(A) = P(\{2\} \cup \{4\} \cup \{6\}) = P(\{2\}) + P(\{4\}) + P(\{6\}) = 1/6 + 1/6 + 1/6 = 1/2.$$

Παρομοίως υπολογίζουμε την πιθανότητα του μονού αποτελέσματος:

$$P(B) = P(\{1\} \cup \{3\} \cup \{5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = 1/6 + 1/6 + 1/6 = 1/2.$$

Παρατήρηση: Στο εξής, για τα στοιχειώδη ενδεχόμενα $\{\omega\}$ θα γράφουμε και $P(\omega)$ εκτός από $P(\{\omega\})$.

Παράδειγμα 1.15. *(Άδικο ζάρι)* Η άνω επιλογή για τις πιθανότητες των στοιχειωδών ενδεχομένων είναι ίσως αυτή που αντιστοιχεί πιο κοντά στην πραγματικότητα, αλλά δεν είναι η μόνη που είναι αποδεκτή από την Θεωρία Πιθανοτήτων. Μια άλλη είναι να επιλέξουμε, για τις πιθανότητες των στοιχειωδών ενδεχόμενων, τις ακόλουθες πιθανότητες:

$$P(1) = P(3) = P(5) = 0, \quad P(2) = P(4) = P(6) = 1/3.$$

Επίσης, για όλα τα άλλα ενδεχόμενα επιλέγουμε την πιθανότητα που προκύπτει με εφαρμογή του τρίτου αξιώματος.

Παράδειγμα 1.16. *(Δύο ζαριές)* Ρίχνουμε ένα ζάρι 2 φορές, οπότε ο δειγματικός χώρος Ω αποτελείται από τα 36 δυνατά αποτελέσματα:

$$\begin{aligned} \Omega = \{&11, 12, 13, 14, 15, 16, \\ &21, 22, \ldots\ldots, 26, \\ &\vdots \quad \vdots \qquad\qquad \vdots \\ &61, 62, \ldots\ldots, 66\}. \end{aligned}$$

1. Όπως και στο Παράδειγμα 1.14, αν η ζαριά είναι δίκαιη λογικά υποθέτουμε ότι το καθένα από τα 36 αποτελέσματα έχει την ίδια πιθανότητα, δηλαδή 1/36. Θα υπολογίσουμε την πιθανότητα των εξής ενδεχομένων:

$$\begin{aligned} A &= \text{«ασσόδυο»} = \{12, 21\}, \\ B &= \text{«εξάρες»} = \{66\}, \\ C &= \text{«6 την πρώτη φορά»} = \{61, 62, 63, 64, 65, 66\}, \\ D &= \text{«άθροισμα 6»} = \{15, 24, 33, 42, 51\}. \end{aligned}$$

Για το A έχουμε, από τις πιο πάνω υποθέσεις,

$$\begin{aligned} P(A) &= P(\{12, 21\}) = P\big(\{12\} \cup \{21\}\big) = P(12) + P(21) \\ &= \frac{1}{36} + \frac{1}{36} = \frac{1}{18} \simeq 0.0555, \end{aligned}$$

όπου και πάλι χρησιμοποιήσαμε το γεγονός ότι τα στοιχειώδη αποτελέσματα είναι πάντοτε ξένα μεταξύ τους. Για το B απλώς έχουμε $P(B) = P(66) = 1/36$, άρα υπάρχει διπλάσια πιθανότητα να φέρουμε ασσόδυο απ' το να φέρουμε εξάρες.

Με την ίδια λογική, για το C έχουμε

$$\begin{aligned} P(C) &= P\big(\{61\} \cup \{62\} \cup \{63\} \cup \{64\} \cup \{65\} \cup \{66\}\big) \\ &= P(61) + P(62) + P(63) + P(64) + P(65) + P(66) \\ &= 6/36 = 1/6, \end{aligned}$$

δηλαδή μόλις αποδείξαμε το διαισθητικά προφανές – ότι η πιθανότητα του να φέρουμε 6 την πρώτη φορά είναι 1/6. Και ακολουθώντας πάλι την ίδια λογική, εύκολα υπολογίζουμε ότι, εφόσον το D αποτελείται από 5 στοιχεία και όλα τα αποτελέσματα είναι ισοπίθανα, $P(D) = 5/36$.

2. Αν η ζαριά δεν είναι δίκαιη, τότε θα πρέπει να δώσουμε σε κάθε ένα από τα αποτελέσματα διαφορετική βαρύτητα, ανάλογα με την διαίσθησή μας για το τι είναι λογικό στο συγκεκριμένο πρόβλημα. Για παράδειγμα, μπορούμε να υποθέσουμε πως

$$\begin{aligned} P(11) = P(22) = \cdots = P(66) &= \frac{1}{18}, \\ P(ij) &= \frac{1}{45},\ i \neq j, \end{aligned}$$

υπολογίζοντας τις πιθανότητες όλων των άλλων ενδεχόμενων μέσω του τρίτου αξιώματος. Η συγκεκριμένη επιλογή θα ήταν κατάλληλη αν, για παράδειγμα, μοντελοποιούσαμε τα αποτελέσματα των ζαριών ενός παίκτη που «τσιμπάει» τα ζάρια και φέρνει διπλές πιο συχνά από άλλες ζαριές.

Παρατήρηση: Τα Παραδείγματα 1.14, 1.15, 1.16 γενικεύονται στα ακόλουθα 2 λήμματα. Το πρώτο από τα λήμματα έχει εφαρμογή σε πειράματα με πεπερασμένο πλήθος αποτελεσμάτων που θέλουμε να μοντελοποιήσουμε ως ισοπίθανα. Το δεύτερο έχει εφαρμογή σε πειράματα με πεπερασμένο πλήθος αποτελεσμάτων, αλλά όπου κάποια αποτελέσματα πρέπει να μοντελοποιηθούν ως πιο πιθανά από άλλα.

Λήμμα 1.2. (Ομοιόμορφο μέτρο πιθανότητας σε πεπερασμένους δειγματικούς χώρους) *Έστω δειγματικός χώρος* $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ *με* $|\Omega| = n$ *αποτελέσματα. Αν θέσουμε*

$$P(A) = \frac{|A|}{|\Omega|}, \quad \forall A \subseteq \Omega,$$

και, συνεπώς,

$$P(\omega_i) = \frac{1}{|\Omega|}, \quad \forall i = 1, 2, \ldots, n,$$

τότε η συνάρτηση P *όπως έχει οριστεί είναι μέτρο πιθανότητας, δηλαδή ικανοποιεί τα αξιώματα των πιθανοτήτων.*

Απόδειξη. Καταρχήν, είναι προφανές ότι όλες οι πιθανότητες, όπως ορίστηκαν, είναι μη αρνητικές. Άρα ικανοποιείται το πρώτο αξίωμα. Επιπλέον, ειδικά για το Ω, έχουμε $P(\Omega) = \frac{|\Omega|}{|\Omega|} = 1$. Άρα ικανοποιείται και το δεύτερο αξίωμα. Έστω τώρα δύο ξένα ενδεχόμενα A_1, A_2. Αφού είναι ξένα, δηλαδή δεν έχουν κοινά στοιχεία, θα πρέπει $|A_1 \cup A_2| = |A_1| + |A_2|$. Άρα:

$$P(A_1 \cup A_2) = \frac{|A_1 \cup A_2|}{|\Omega|} = \frac{|A_1|}{|\Omega|} + \frac{|A_2|}{|\Omega|} = P(A_1) + P(A_2),$$

και επαληθεύσαμε και το τρίτο αξίωμα για την περίπτωση δύο ενδεχόμενων. Η γενίκευση στην περίπτωση περισσότερων ενδεχόμενων είναι απλή. □

Λήμμα 1.3. (Αυθαίρετο μέτρο πιθανότητας σε πεπερασμένους δειγματικούς χώρους) *Έστω δειγματικός χώρος* $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ *με* $|\Omega| = n$ *αποτελέσματα. Έστω οποιοδήποτε σύνολο αριθμών* p_i, $i = 1, \ldots, n$, *όπου*

$$p_i \geq 0, \; i = 1, \ldots, n, \quad \sum_{i=1}^{n} p_i = 1.$$

Αν θέσουμε

$$P(A) = \sum_{i:\ \omega_i \in A} p_i,$$

και συνεπώς

$$P(\omega_i) = p_i, \quad i = 1, \ldots, n,$$

τότε η συνάρτηση P όπως έχει οριστεί είναι μέτρο πιθανότητας, δηλαδή ικανοποιεί τα αξιώματα των πιθανοτήτων.

Απόδειξη. Καταρχήν, είναι προφανές ότι όλες οι πιθανότητες είναι μη αρνητικές. Άρα ικανοποιείται το πρώτο αξίωμα. Επιπλέον, ειδικά για το Ω, έχουμε

$$P(\Omega) = \sum_{i:\ \omega_i \in \Omega} p_i = \sum_{i=1}^{n} p_i = 1.$$

Άρα ικανοποιείται και το δεύτερο αξίωμα. Έστω τώρα δύο ξένα ενδεχόμενα A_1, A_2. Θα έχουμε:

$$P(A_1 \cup A_2) = \sum_{i:\ \omega_i \in A_1 \cup A_2} p_i = \sum_{i:\ \omega_i \in A_1} p_i + \sum_{i:\ \omega_i \in A_2} p_i = P(A_1) + P(A_2).$$

και επαληθεύσαμε και το τρίτο αξίωμα για την περίπτωση δύο ενδεχόμενων. (Παρατηρήστε ότι στη δεύτερη ισότητα χρησιμοποιήσαμε το ότι τα A_1, A_2 είναι ξένα.) Η γενίκευση στην περίπτωση περισσότερων ενδεχόμενων είναι απλή. □

Παράδειγμα 1.17. *(Περίεργα ζάρια)* Έστω ένα (όχι απαραίτητα δίκαιο) ζάρι για το οποίο γνωρίζουμε ότι η πιθανότητα να έρθει 1 ή 2 είναι 1/3, ενώ η πιθανότητα να έρθει 2 ή 3 είναι επίσης 1/3. Ποια είναι η μέγιστη δυνατή και η ελάχιστη δυνατή τιμή για την πιθανότητα να έρθει 2, λαμβάνοντας υπόψιν όλα τα δυνατά μέτρα πιθανότητας που υπάρχουν;

Για να απαντήσουμε το ερώτημα, έστω ο δειγματικός χώρος $\Omega = \{1, 2, 3, 4, 5, 6\}$, και έστω P το μέτρο πιθανότητας που περιγράφει τις πιθανότητες αυτού του ζαριού. Μας έχει δοθεί ότι:

$$P(\{1, 2\}) = P(1) + P(2) = 1/3, \quad P(\{2, 3\}) = P(2) + P(3) = 1/3.$$

Για το κάτω φράγμα, προφανώς έχουμε $P(2) \geq 0$, και η τιμή $P(2) = 0$ είναι εφικτή αν επιλέξουμε, για παράδειγμα, ως μέτρο πιθανότητας το

$$P(1) = P(3) = P(5) = 1/3, \quad P(2) = P(4) = P(6) = 0,$$

που είναι συμβατό με τα δεδομένα του προβλήματος.

Για το άνω φράγμα, παρατηρούμε ότι θα πρέπει η πιθανότητα του 2 είναι το πολύ ίση με την πιθανότητα του ενδεχόμενου να έρθει 1 ή 2, που είναι 1/3. (Από ποια εξίσωση

προκύπτει αυτό;) Αυτό το άνω φράγμα είναι εφικτό και επιτυγχάνεται, για παράδειγμα, για το ακόλουθο μέτρο πιθανότητας:

$$P(1) = P(3) = 0, \quad P(2) = P(4) = 1/3, \quad P(5) = P(6) = 1/6.$$

Παρατηρήστε ότι τα πιο πάνω μέτρα πιθανότητας ικανοποιούν όλα τα αξιώματα πιθανοτήτων, και επομένως είναι «αποδεκτά». Το κατά πόσο θα είναι χρήσιμα, εξαρτάται από το πόσο ανταποκρίνονται στην πραγματικότητα, και μπορούμε συνεπώς να τα χρησιμοποιήσουμε ως μοντέλα. (Από αυτή την άποψη, στις περισσότερες περιπτώσεις δεν είναι αποδεκτά, καθώς τα περισσότερα ζάρια ανταποκρίνονται στις συνθήκες $P(i) \simeq 1/6$ για κάθε $1 \leq i \leq 6$.)

Παρατήρηση: Το επόμενο παράδειγμα δείχνει πως θα μπορούσαμε να ορίσουμε ένα μέτρο πιθανότητας σε ένα μη πεπερασμένο δειγματικό χώρο.

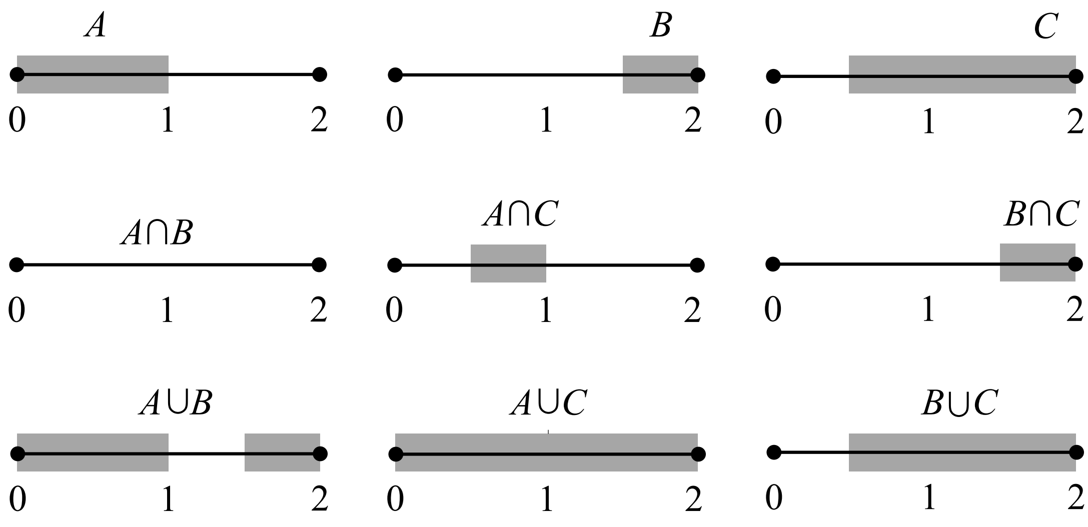
Παράδειγμα 1.18. *(Τυχαία συνάντηση)* Ο Σταύρος και ο Γιάννης έχουν ορίσει να συναντηθούν σε ένα μπαρ. Ο Σταύρος έχει έρθει στην ώρα του, αλλά για τον Γιάννη γνωρίζουμε ότι μπορεί να εμφανιστεί οποιαδήποτε στιγμή μέσα στις επόμενες δύο ώρες, χωρίς προτίμηση σε κάποια στιγμή ή διάστημα. Θεωρούμε τυχαίο πείραμα την άφιξη του Γιάννη, και ορίζουμε ως δειγματικό χώρο το διάστημα $\Omega = [0, 2]$. Ορίζουμε τα ενδεχόμενα $A = [0, 1]$ ο Γιάννης να έρθει εντός της πρώτης ώρας, $B = [3/2, 2]$ ο Γιάννης να έρθει εντός του τελευταίου μισάωρου, και $C = [1/2, 2]$ ο Γιάννης να έρθει με τουλάχιστον μισή ώρα καθυστέρηση. Θα υπολογίσουμε τις πιθανότητες των ενδεχόμενων A, B, C, $A \cap B$, $A \cap C$, $B \cap C$, $A \cup B$, $A \cup C$, $B \cup C$.

Επειδή ο Γιάννης δεν έχει προτίμηση σε κάποιο διάστημα, είναι λογικό να υποθέσουμε ότι η πιθανότητα ενός οποιουδήποτε διαστήματος E είναι ανάλογη του μήκους του, $\ell(E)$. Προκειμένου να είναι η πιθανότητα όλου του δειγματικού χώρου Ω ίση με τη μονάδα, προκύπτει ότι πρέπει $P(E) = \ell(E)/2$. Αν θέλουμε να υπολογίσουμε την πιθανότητα οποιουδήποτε άλλου συνόλου, το γράφουμε ως ένωση ξένων διαστημάτων, και εφαρμόζουμε το τρίτο αξίωμα. Έχουμε ότι

$$\begin{aligned} A = [0,1] &\Rightarrow P(A) = 1/2, \\ B = [3/2,2] &\Rightarrow P(B) = 1/4, \\ C = [1/2,2] &\Rightarrow P(C) = 3/4, \end{aligned}$$

Επιπλέον,

$$\begin{aligned} A \cap B = \emptyset &\Rightarrow P(A \cap B) = 0, \\ A \cap C = [1/2,1] &\Rightarrow P(A \cap C) = 1/4, \\ B \cap C = [3/2,2] &\Rightarrow P(B \cap C) = 1/4. \end{aligned}$$



Σχήμα 1.6: Παράδειγμα 1.18.

Τέλος,

$$
\begin{aligned}
P(A \cup B) &= P([0,1] \cup [3/2,2]) \;=\; P([0,1]) + P([3/2,2]) \;=\; 1/2 + 1/4 \;=\; 3/4, \\
P(A \cup C) &= P([0,1] \cup [1/2,2]) \;=\; P([0,2]) \;=\; 1. \\
P(B \cup C) &= P([3/2,2] \cup [1/2,2]) \;=\; P([1/2,2]) = 3/4.
\end{aligned}
$$

Όλα τα πιο πάνω ενδεχόμενα εμφανίζονται στο Σχήμα 1.6.

1.4 Ιδιότητες Πιθανοτήτων

Ξεκινώντας από τα τρία αξιώματα των πιθανοτήτων μπορούμε να αποδείξουμε μια σειρά από ιδιότητες, όλες διαισθητικά αναμενόμενες.

Λήμμα 1.4. (Ιδιότητες πιθανοτήτων) *Έστω Ω ένας δειγματικός χώρος και P ένα μέτρο πιθανότητας. Στα ακόλουθα, A, B, C είναι ενδεχόμενα του Ω. Ισχύουν οι ακόλουθες ιδιότητες:*

1. $P(A') = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. $P(A) \leq 1$.
4. $A \subseteq B \Rightarrow P(A) \leq P(B)$.
5. $P(A \cap B') = P(A) - P(A \cap B)$.
6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
7. $P(A \cup B) \leq P(A) + P(B)$. *(Γνωστή ως Φράγμα της Ένωσης)*
8. $P(A \cup B \cup C) = P(A) + P(A' \cap B) + P(A' \cap B' \cap C)$.
9. $$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Απόδειξη. 1.

$$1 = P(\Omega) = P(A \cup A') = P(A) + P(A') \Rightarrow P(A') = 1 - P(A).$$

2.

$$P(\emptyset) = 1 - P(\emptyset') = 1 - P(\Omega) = 0.$$

3. Προκύπτει από το πρώτο σκέλος, αφού η $P(A')$ είναι μη αρνητική (λόγω του πρώτου αξιώματος).

4. Παρατηρούμε πως $B = A \cup (B \cap A')$. Πράγματι, εύκολα μπορούμε να δείξουμε πως το ένα σύνολο είναι υποσύνολο του άλλου και αντιστρόφως. Επιπλέον, τα A, $B \cap A'$ είναι ξένα μεταξύ τους (εξετάστε την τομή τους για να βεβαιωθείτε). Άρα:

$$P(B) = P(A \cup (B \cap A')) = P(A) + P(B \cap A').$$

Η δεύτερη ισότητα προκύπτει από το τρίτο αξίωμα. Επειδή $P(B \cap A') \geq 0$ (από το πρώτο αξίωμα), προκύπτει το ζητούμενο.

5. Αρκεί να παρατηρήσουμε πως $A = (A \cap B) \cup (A \cap B')$, και πως τα $A \cap B$, $A \cap B'$ είναι ξένα.

6. Παρατηρούμε πως $A \cup B = A \cup (A' \cap B)$. Επιπλέον, τα A, $A' \cap B$ είναι ξένα, άρα θα ισχύει

$$P(A \cup B) = P(A \cup (A' \cap B)) = P(A) + P(A' \cap B).$$

   Επιπλέον, από το προηγούμενο σκέλος,

$$P(A' \cap B) = P(B) - P(A \cap B).$$

   Συνδυάζοντας τις άνω δύο εξισώσεις προκύπτει το ζητούμενο.

7. Προκύπτει άμεσα από την προηγούμενη.

8. Αρκεί να παρατηρήσουμε πως $A \cup B \cup C = A \cup (A' \cap B) \cup (A' \cap B' \cap C)$, και πως τα A, $A' \cap B$, $A' \cap B' \cap C$ είναι ξένα.

9. Παρατηρούμε πως:

$$\begin{aligned}
P(A \cup B \cup C) &= P(A \cup (B \cup C)) \\
&= P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\
&= P(A) + P(B) + P(C) - P(B \cap C) \\
&\quad -P((A \cap B) \cup (A \cap C)) \\
&= P(A) + P(B) + P(C) - P(B \cap C) \\
&\quad -P(A \cap B) - P(A \cap C) + P((A \cap B) \cap (A \cap C)) \\
&= P(A) + P(B) + P(C) - P(A \cap B) \\
&\quad -P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).
\end{aligned}$$

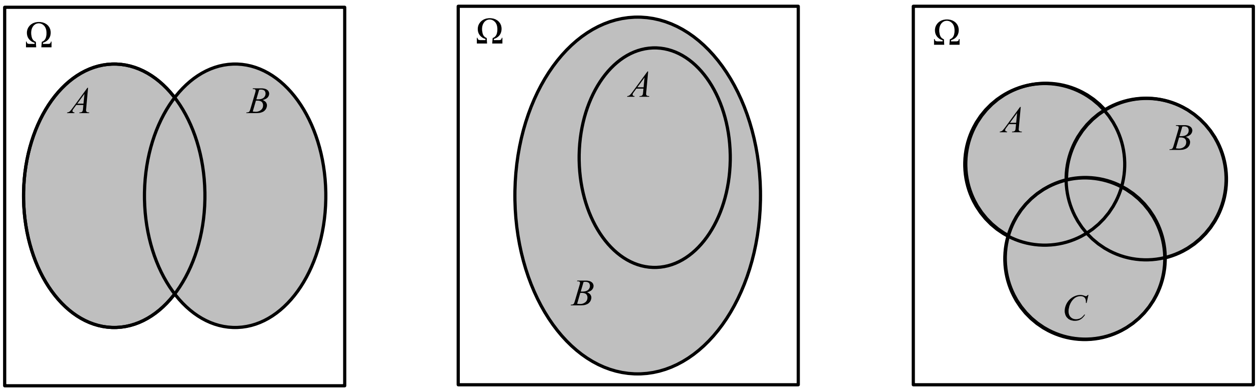
□

Παρατηρήσεις

1. Βεβαιωθείτε ότι μπορείτε να δικαιολογήσετε όλα τα βήματα των άνω αποδείξεων.

2. Βεβαιωθείτε ότι μπορείτε να δώσετε μια διαισθητική εξήγηση για όλες τις ιδιότητες.

3. Ορισμένες εκ των άνω γενικεύονται και για την περίπτωση n ενδεχόμενων. Μπορείτε να κάνετε τις γενικεύσεις;



Σχήμα 1.7: Μερικά διαγράμματα Venn για το Λήμμα 1.4.

4. Όλες οι άνω ιδιότητες, μπορούν αν περιγραφούν και με διαγράμματα Venn, τα οποία σε πολλές περιπτώσεις δείχνουν το δρόμο για την αυστηρή απόδειξη. Επίκληση όμως σε διάγραμμα Venn ΔΕΝ είναι αυστηρή απόδειξη. Ορισμένα διαγράμματα Venn που αντιστοιχούν σε κάποιες από τις άνω ιδιότητες υπάρχουν στο Σχήμα 1.7. Μπορείτε να βρείτε ποιο διάγραμμα αντιστοιχεί πού; Μπορείτε να φτιάξετε διαγράμματα για τις υπόλοιπες ιδιότητες;

5. Το σύνολο $A \cap B'$ περιλαμβάνει όλα τα στοιχεία του A αν αφαιρέσουμε τα στοιχεία του B, γι΄ αυτό συχνά καλείται η <u>διαφορά</u> του B από το A, και συμβολίζεται $A - B$.

6. Στα ακόλουθα, προκειμένου να έχουμε πιο συνοπτικούς τύπους, μερικές φορές θα παραλείπουμε να γράφουμε το σύμβολο της τομής, και θα το υπονοούμε. Δηλαδή: $A \cap B = AB$. Παρατηρήστε ότι κάτι ανάλογο γίνεται και με την πράξη του πολλαπλασιασμού. Για παράδειγμα, έχουμε την ακόλουθη ιδιότητα:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

Παράδειγμα 1.19. *(Ένας δειγματικός χώρος με 3 στοιχεία)* Ένα τυχαίο πείραμα έχει δειγματικό χώρο $\Omega = \{a, b, c\}$. Έστω πως κάποιο μέτρο πιθανότητας P ικανοποιεί τις σχέσεις, $P(\{a, c\}) = 9/16$, $P(\{a, b\}) = 3/4$. Θα υπολογίσουμε τις πιθανότητες όλων των στοιχειωδών ενδεχομένων. Τα αποτελέσματα είναι ξένα μεταξύ τους, συνεπώς η πιθανότητα της ένωσής τους ισούται με το άθροισμα των πιθανοτήτων τους:

$$P(\{a, c\}) = P(\{a\} \cup \{c\}) = \frac{9}{16} \Rightarrow P(a) + P(c) = \frac{9}{16}.$$

Ομοίως προκύπτει ότι:

$$P(a) + P(b) = \frac{3}{4}.$$

Τέλος, ο δειγματικός χώρος Ω έχει $P(\Omega) = 1$, συνεπώς έχουμε:

$$P(a) + P(b) + P(c) = P(\{a\} \cup \{b\} \cup \{c\}) = P(\Omega) = 1.$$

Οι τρεις άνω εξισώσεις σχηματίζουν ένα γραμμικό σύστημα εξισώσεων 3×3 ως προς $P(a), P(b), P(c)$, με μοναδική λύση:

$$P(a) = \frac{5}{16}, \quad P(b) = \frac{7}{16}, \quad P(c) = \frac{1}{4}.$$

Παράδειγμα 1.20. *(Τι λένε οι πιθανοθεωρίστες στα παιδιά τους)* Στην Αττική, το 95% των εγκλημάτων συμβαίνει τη νύχτα, και το 54% συμβαίνει μέσα στην Αθήνα. Αν μόνο 2% των εγκλημάτων συμβαίνουν μέρα στην Αθήνα, τι ποσοστό συμβαίνει νύχτα στην Αθήνα; Τι ποσοστό συμβαίνει νύχτα έξω από την Αθήνα;

Για να απαντήσουμε τα ερωτήματα, ορίζουμε ως δειγματικό χώρο Ω το σύνολο όλων των εγκλημάτων που μπορούν να συμβούν στην Αττική, ως A το ενδεχόμενο ένα έγκλημα να έγινε στην Αθήνα, και ως N το ενδεχόμενο το έγκλημα να έγινε νύχτα. Μας δίνεται ότι

$$P(N) = 0.95, \quad P(A) = 0.54, \quad P(N' \cap A) = 0.02,$$

και ζητούνται οι πιθανότητες $P(N \cap A)$ και $P(N \cap A')$.

Παρατηρούμε ότι

$$P(N \cap A) = P(A) - P(N' \cap A) = 0.52.$$

Ομοίως,

$$P(N \cap A') = P(N) - P(N \cap A) = 0.95 - 0.52 = 0.43.$$

Μπορείτε να βρείτε ποια (ή ποιες) από τις ιδιότητες του Λήμματος 1.4 χρησιμοποιήσαμε;

Παράδειγμα 1.21. *(Πιθανότητα για ακριβώς ένα ενδεχόμενο)* Έστω A και B δύο ενδεχόμενα. Έστω το ενδεχόμενο C να πραγματοποιείται μόνο ένα εκ των A, B. Θα υπολογίσουμε την πιθανότητά του. Παρατηρούμε καταρχήν πως $C = (A \cap B') \cup (A' \cap B)$. Όμως, $C \cup (A \cap B) = A \cup B$, και $C \cap (A \cap B) = \emptyset$. Συνεπώς,

$$P(A \cup B) = P(C \cup (A \cap B)) = P(C) + P(A \cap B).$$

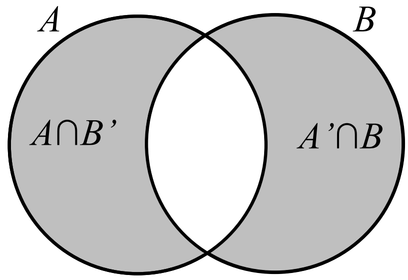
Επιπλέον, έχουμε τη γνωστή ισότητα

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Συνδυάζοντας τις δύο άνω εξισώσεις, προκύπτει τελικά πως

$$P(C) = P(A) + P(B) - 2P(A \cap B).$$

Το αποτέλεσμα εύκολα μπορεί να αιτιολογηθεί και διαισθητικά, με χρήση διαγράμματος Venn. Δείτε το Σχήμα 1.8.



Σχήμα 1.8: Παράδειγμα 1.21.

Παράδειγμα 1.22. *(Φράγμα της Ένωσης)* Θα δείξουμε το *φράγμα της ένωσης (union bound)* για την περίπτωση n συνόλων $A_1, A_2, \ldots, A_n$:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

Έστω τα ενδεχόμενα

$$B_i = A_i \cap \left(\bigcap_{k=1}^{i-1} A'_k\right).$$

Παρατηρούμε ότι το ενδεχόμενο B_i περιλαμβάνει όλα τα αποτελέσματα που περιέχει το A_i αλλά κανένα A_k με $k < i$. Επιπλέον, παρατηρούμε ότι τα B_i είναι ξένα μεταξύ τους. Πράγματι, αν $i < j$, από τον τρόπο που ορίστηκαν προκύπτει ότι το B_i είναι υποσύνολο του A_i, ενώ το B_j είναι υποσύνολο του A'_i. Από την άλλη,

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i.$$

Πράγματι, έστω $x \in \cup_{i=1}^{n} B_i$. Τότε $x \in B_i$ για κάποιο i, άρα από τον ορισμό του B_i θα έχουμε $x \in A_i$, άρα $x \in \cup_{i=1}^{n} A_i$. Έστω από την άλλη πως $x \in \bigcup_{i=1}^{n} A_i$. Τότε $x \in A_i$ για ορισμένα i. Ας πάρουμε το μικρότερο από αυτά. Τότε, από την κατασκευή του B_i, θα έχουμε $x \in B_i$, άρα και $x \in \bigcup_{i=1}^{n} B_i$.

Τελικά προκύπτει:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=1}^{n} B_i\right) = \sum_{i=1}^{n} P(B_i) \leq \sum_{i=1}^{n} P(A_i),$$

και αποδείχτηκε το ζητούμενο. Η δεύτερη ισότητα προκύπτει από το γεγονός ότι τα B_i είναι ξένα, και η ανισότητα από το ότι $B_i \subseteq A_i \Rightarrow P(B_i) \leq P(A_i)$.

Μπορείτε να αποδείξετε το φράγμα της ένωσης χρησιμοποιώντας, αντί για την άνω μέθοδο, επαγωγή;

Κεφάλαιο 2

Συνδυαστική

Όπως είδαμε σε κάποια παραδείγματα του προηγούμενου κεφαλαίου, συχνά συναντάμε περιπτώσεις όπου τα αποτελέσματα ενός τυχαίου πειράματος με δειγματικό χώρο Ω έχουν πεπερασμένο πλήθος $|\Omega|$ και θέλουμε να τα μοντελοποιήσουμε ως ισοπίθανα. Σε αυτές τις περιπτώσεις μπορεί να εφαρμοστεί το Λήμμα 1.2, σύμφωνα με το οποίο, σε ένα τέτοιο τυχαίο πείραμα, η πιθανότητα ενός ενδεχόμενου A που περιλαμβάνει $|A|$ αποτελέσματα είναι

$$P(A) = \frac{|A|}{|\Omega|}.$$

Παράδειγμα 2.1. Ρίχνουμε διαδοχικά δύο ζάρια. Ο δειγματικός χώρος περιλαμβάνει τα ακόλουθα 36 αποτελέσματα, που θεωρούμε ισοπίθανα:

$$\begin{aligned} \Omega \;=\; \{&11, 12, 13, 14, 15, 16,\\ &21, 22, \;\ldots\ldots\;, 26,\\ &\;\vdots \quad \vdots \qquad\qquad\quad \vdots\\ &61, 62, \;\ldots\ldots\;, 66\}. \end{aligned}$$

Ποια είναι η πιθανότητα του ενδεχόμενου A να φέρουμε τουλάχιστον ένα 5; Παρατηρούμε πως

$$A = \{15,\ 25,\ 35,\ 45,\ 55,\ 65,\ 51,\ 52,\ 53,\ 54,\ 56\} \Rightarrow |A| = 11 \Rightarrow P(A) = \frac{|A|}{|\Omega|} = \frac{11}{36}.$$

Στο άνω παράδειγμα, ήταν απλό να υπολογίσουμε το μέγεθος των εμπλεκόμενων συνόλων. Γενικά όμως, η απαρίθμηση των στοιχείων ενός συνόλου δεν είναι προφανής. Η Συνδυαστική είναι ο μαθηματικός τομέας που μας προσφέρει ακριβώς τα εργαλεία που χρειαζόμαστε για αυτούς τους υπολογισμούς. Πιο κάτω θα δούμε μια σειρά από σχετικά απλά αποτελέσματα της Συνδυαστικής, και μέσα από παραδείγματα θα δείξουμε με ποιους τρόπους αυτά τα αποτελέσματα χρησιμοποιούνται για την απάντηση ερωτημάτων σε προβλήματα πιθανοτήτων. Σημείο εκκίνησης είναι η ακόλουθη απλή παρατήρηση:

Λήμμα 2.1. (Πολλαπλασιαστική Αρχή) *Όταν εκτελούνται διαδοχικά δύο πειράματα, εκ των οποίον το πρώτο έχει N δυνατά αποτελέσματα και το δεύτερο έχει M δυνατά αποτελέσματα, τότε το νέο, σύνθετο πείραμα έχει $M \times N$ δυνατά αποτελέσματα. Και πιο γενικά, αν εκτελεστούν διαδοχικά k πειράματα, καθένα με $N_1, N_2, \ldots, N_k$ δυνατά αποτελέσματα, τότε το νέο, σύνθετο πείραμα έχει $N_1 \times N_2 \times \ldots \times N_k$ δυνατά αποτελέσματα.*

Παράδειγμα 2.2.

1. Εφόσον η ρίψη ενός ζαριού έχει 6 δυνατά αποτελέσματα, οι δύο διαδοχικές ρίψεις έχουν $6 \times 6 = 36$ δυνατά αποτελέσματα, οι τρεις διαδοχικές ρίψεις έχουν $6 \times 6 \times 6 = 216$ δυνατά αποτελέσματα, και γενικά οι k διαδοχικές ρίψεις έχουν 6^k δυνατά αποτελέσματα.

2. Επιλέγουμε έναν από τους 5 υπολογιστές ενός εργαστηρίου (5 δυνατά αποτελέσματα) και αποφασίζουμε να του εγκαταστήσουμε λειτουργικό σύστημα Windows ή Linux (2 δυνατά αποτελέσματα). Συνολικά υπάρχουν $5 \times 2 = 10$ δυνατά αποτελέσματα.

3. Βάζουμε σε σειρά 4 βιβλία. Έχουμε 4 επιλογές για το πρώτο βιβλίο, 3 επιλογές για το δεύτερο, 2 για το τρίτο, και 1 μόνο για το τελευταίο. Άρα συνολικά έχουμε $4 \times 3 \times 2 = 24$ τρόπους για να βάλουμε 4 βιβλία στη σειρά. Παρατηρήστε ότι σε αυτή την περίπτωση, κάθε πείραμα που διεξάγεται επηρεάζει τα πειράματα που θα ακολουθήσουν (και για την ακρίβεια, ποια είναι τα δυνατά τους αποτελέσματα).

2.1 Διατάξεις

Ορισμός 2.1. (Διατάξεις και μεταθέσεις) *Έστω ένα σύνολο X που αποτελείται από πεπερασμένο πλήθος στοιχείων $|X| = N$. Καλούμε διάταξη κάθε διατεταγμένη k-άδα στοιχείων του X, διακριτών μεταξύ τους. Πρέπει $k \leq N$. Καλούμε μετάθεση κάθε διατεταγμένη N-άδα στοιχείων.*

Παράδειγμα 2.3. Σε ένα αγώνα που συμμετέχουν 100 αθλητές, πόσες δυνατές διατάξεις υπάρχουν για τις πρώτες τρεις θέσεις; Μπορούμε να φανταστούμε ότι εκτελούμε τρία διαδοχικά πειράματα. Το πρώτο πείραμα, που είναι ο καθορισμός της πρώτης θέσης, έχει 100 αποτελέσματα. Το δεύτερο πείραμα έχει 99 αποτελέσματα (αφού έχουν μείνει πλέον 99 αθλητές να διαγωνίζονται για την δεύτερη θέση). Το τρίτο έχει 98 αποτελέσματα. Άρα έχουμε $100 \times 99 \times 98 = 970200$ διατάξεις.

Επίσης, πόσες δυνατές κατατάξεις και για τους 100 αθλητές, δηλαδή μεταθέσεις, υπάρχουν; Θα εκτελέσουμε τώρα 100 διαδοχικά πειράματα. Για την πρώτη θέση έχουμε 100 επιλογές. Για την δεύτερη θέση έχουμε 99 επιλογές, κ.ο.κ. Έτσι, εφαρμόζοντας την πολλαπλασιαστική αρχή, έχουμε

$$100 \times 99 \times 98 \times \cdots \times 3 \times 2 \times 1 = 100!$$

δυνατές μεταθέσεις, δηλαδή 100! δυνατούς τρόπους που μπορούν να διαταχθούν 100 άτομα.

Παρατήρηση: Θυμίζουμε πως για κάθε ακέραιο αριθμό $N \geq 1$ το «N παραγοντικό» συμβολίζεται ως $N!$ και ορίζεται ως το γινόμενο $N! = N(N-1)\dots 2 \times 1$. Επίσης, για λόγους ευκολίας ορίζουμε συμβατικά το $0! = 1$.

Λήμμα 2.2. (Πλήθος διατάξεων/μεταθέσεων) *Το πλήθος όλων των δυνατών διατάξεων $k \leq N$ αντικειμένων που επιλέγονται από N αντικείμενα ισούται με:*

$$(N)_k \triangleq N(N-1)\dots(N-k+1) = \frac{N!}{(N-k)!}.$$

Συνεπώς, υπάρχουν $N!$ δυνατές μεταθέσεις N αντικειμένων.

Απόδειξη. Μπορούμε να φανταστούμε την επιλογή της διάταξης ως την εκτέλεση διαδοχικών πειραμάτων. Έχουμε N επιλογές για το πρώτο πείραμα που είναι η επιλογή του πρώτου αντικειμένου, $(N-1)$ για το δεύτερο πείραμα, κ.ο.κ., μέχρι το πείραμα k, για το οποίο έχουμε $(N-k+1)$ επιλογές. Άρα από την πολλαπλασιαστική αρχή, βρίσκουμε

πως το πλήθος των διατάξεων είναι

$$N(N-1)\dots(N-k+2)(N-k+1) = \frac{N(N-1)\dots(N-k+1)(N-k)!}{(N-k)!} = \frac{N!}{(N-k)!}.$$

Όταν $k = N$, ο άνω τύπος δίνει $N!$ μεταθέσεις. □

Παράδειγμα 2.4. Έστω ότι έχουμε μια συνηθισμένη τράπουλα 52 φύλλων. Από το Λήμμα 2.2 υπάρχουν $52! \simeq 8 \times 10^{67}$ δυνατές μεταθέσεις για τα φύλλα της τράπουλας!

Αν επιλέξουμε 3 φύλλα στην τύχη, πόσες δυνατές (διατεταγμένες) τριάδες υπάρχουν; Από το ίδιο λήμμα, το πλήθος τους είναι

$$|\Omega| = \frac{52!}{(52-3)!} = \frac{52 \times 51 \times 50 \times 49!}{49!} = 52 \times 51 \times 50 = 132600.$$

Αν υποθέσουμε τώρα ότι η επιλογή των τριών φύλλων είναι εντελώς τυχαία, ή, πιο αυστηρά, ότι όλα τα αποτελέσματα του πειράματος της διαδοχικής επιλογής τριών φύλλων είναι ισοπίθανα, ποια είναι η πιθανότητα του ενδεχομένου A να επιλέξουμε τρεις άσσους; Εφαρμόζοντας το Λήμμα 1.2, η πιθανότητα είναι

$$P(A) = \frac{|A|}{|\Omega|}.$$

Πρέπει λοιπόν να υπολογίσουμε το πλήθος $|A|$, αφού το $|\Omega|$ μας είναι γνωστό. Φανταστείτε ότι εκτελείται ένα νέο πείραμα, στο οποίο επιλέγονται διαδοχικά 3 άσσοι ανάμεσα σε 4. Τα αποτελέσματα αυτού του πειράματος, με χρήση του Λήμματος 2.2, έχουν πλήθος $\frac{4!}{(4-3)!}$. Παρατηρήστε όμως πως τα αποτελέσματα του νέου πειράματος ταυτίζονται με τα αποτελέσματα του αρχικού πειράματος που ανήκουν στο A. Άρα, $|A| = \frac{4!}{(4-3)!} = 4!/1! = 24$, και τελικά

$$P(\text{«επιλέξαμε 3 άσσους»}) = P(A) = \frac{|A|}{|\Omega|} = \frac{24}{132600} = \frac{1}{5525} \simeq 0.02\%.$$

2.2 Συνδυασμοί

Ορισμός 2.2. (Συνδυασμοί) *Έστω ένα σύνολο* X *που αποτελείται από πεπερασμένο πλήθος στοιχείων* $|X| = N$. *Καλούμε συνδυασμό (ή μη-διατεταγμένη επιλογή) κάθε μη διατεταγμένη* k*-άδα στοιχείων του* X*, διακριτών μεταξύ τους. Πρέπει* $k \leq N$.

Παράδειγμα 2.5.
1. Επιλέγουμε 2 από 4 βιβλία για να πάρουμε μαζί μας στις διακοπές. Αν τα βιβλία είναι τα A, B, C, D, τότε οι δυνατοί συνδυασμοί είναι οι

$$AB, \quad AC, \quad AD, \quad BC, \quad BD, \quad CD.$$

2. Ένας λοχίας επιλέγει 4 άτομα από ένα λόχο 70 ατόμων για να κάνουν μια εργασία από κοινού.

3. Επιλέγουμε 5 χαρτιά από μια τράπουλα 52 χαρτιών.

Στις άνω περιπτώσεις, δεν έχει σημασία η σειρά της επιλογής των στοιχείων. Στις τελευταίες δύο περιπτώσεις, το πλήθος των συνδυασμών είναι πολύ μεγάλο, για αυτό και δεν τους απαριθμούμε. Θα ήταν όμως πολύ χρήσιμο να ξέρουμε πόσοι είναι!

Λήμμα 2.3. *Το πλήθος όλων των συνδυασμών* k *αντικειμένων που επιλέγονται από* N *αντικείμενα ισούται με:*

$$\binom{N}{k} \triangleq \frac{N!}{k!(N-k)!}.$$

Το $\binom{N}{k}$ *καλείται διωνυμικός συντελεστής.*

Απόδειξη. Έστω ότι το πλήθος των συνδυασμών είναι y. Από το Λήμμα 2.2, κάθε ομάδα επιλεγμένων στοιχείων μπορεί να μετατεθεί με $k!$ τρόπους. Άρα, το συνολικό πλήθος των *διατεταγμένων* ομάδων είναι $y \times k!$ Αλλά από το Λήμμα 2.2, αυτό ισούται με $N!/(N-k)!$. Άρα έχουμε

$$y \times k! = \frac{N!}{(N-k)!} \Rightarrow y = \frac{N!}{k!(N-k)!}.$$

□

Παράδειγμα 2.6. Επιλέγουμε τυχαία 3 βιβλία από 10, που αποτελούνται από 5 συγγράμματα μαθημάτων και 5 manual υπολογιστών. Ποια η πιθανότητα να είναι όλα manual; Να είναι ακριβώς δύο manual κι ένα σύγγραμμα;

Εφόσον δεν μας ενδιαφέρει η σειρά με την οποία επιλέγονται τα βιβλία, ορίζουμε

$$\begin{aligned} \Omega &= \{\text{όλες οι μη-διατεταγμένες 3άδες βιβλίων}\}, \\ A &= \{\text{οι μη-διατεταγμένες 3άδες από manual}\}, \\ B &= \{\text{οι μη-διατεταγμένες 3άδες με 2 manual και ένα σύγγραμμα}\}. \end{aligned}$$

Θα υπολογίσουμε τα μεγέθη αυτών των συνόλων.

Καταρχήν παρατηρούμε πως το πλήθος των μη διατεταγμένων 3άδων βιβλίων είναι

$$|\Omega| = \binom{10}{3} = 120.$$

Για να προσδιορίσουμε το μέγεθος του A, φανταζόμαστε πως κάνουμε ένα καινούργιο πείραμα, στο οποίο επιλέγουμε 3 από τα 5 manual χωρίς διάταξη. Αυτό το πείραμα έχει πλήθος αποτελεσμάτων $\binom{5}{3}$, και παρατηρήστε πως αυτά τα αποτελέσματα είναι ακριβώς τα αποτελέσματα του αρχικού πειράματος που ανήκουν στο A, άρα

$$|A| = \binom{5}{3} = 10.$$

Για να προσδιορίσουμε το μέγεθος του B, φανταζόμαστε πως κάνουμε ένα καινούργιο, σύνθετο πείραμα, που αποτελείται από δύο υποπειράματα. Στο πρώτο, επιλέγουμε 2 manual ανάμεσα σε 5, ενώ στο δεύτερο επιλέγουμε 1 σύγγραμμα ανάμεσα σε 5. Τα δύο πειράματα έχουν $\binom{5}{2}$ και $\binom{5}{1}$ δυνατά αποτελέσματα αντιστοίχως, ενώ το σύνθετο πείραμα, σύμφωνα με την πολλαπλασιαστική αρχή, έχει πλήθος αποτελεσμάτων ίσο με το γινόμενό τους. Παρατηρήστε πως το σύνθετο αυτό πείραμα έχει αποτελέσματα ακριβώς τα αποτελέσματα του αρχικού πειράματος που ανήκουν στο B. Άρα τελικά

$$|B| = \binom{5}{2} \times \binom{5}{1} = 50.$$

Άρα τελικά, χρησιμοποιώντας το Λήμμα 1.2, έχουμε τις πιθανότητες

$$\begin{aligned} P(\text{«επιλέξαμε 3 manual»}) &= P(A) = \frac{|A|}{|\Omega|} = \frac{10}{120} = \frac{1}{12}, \\ P(\text{«επιλέξαμε 2 manual και ένα σύγγραμμα»}) &= P(B) = \frac{|B|}{|\Omega|} = \frac{50}{120} = \frac{5}{12}. \end{aligned}$$

Παράδειγμα 2.7. Επιλέγοντας 5 άτομα από 100 για να συγκροτήσουμε μια επιτροπή, αν δεν μας ενδιαφέρει η σειρά επιλογής, υπάρχουν

$$\binom{100}{5} = \frac{100!}{5!(100-5)!} = \frac{100 \times 99 \times 98 \times 97 \times 96 \times 95!}{5!95!} = 75287520$$

δυνατές 5άδες που μπορούμε να επιλέξουμε.

Έστω τώρα ότι τα 100 άτομα αποτελούνται από 40 άνδρες και 60 γυναίκες, και ότι η επιλογή μας είναι εντελώς τυχαία. Θα εξετάσουμε τα εξής ερωτήματα:

1. Ποια η πιθανότητα να επιλέξουμε 2 άνδρες και 3 γυναίκες;

2. Ποια η πιθανότητα να επιλέξουμε 4 άνδρες και 1 γυναίκα;

3. Ποια η πιθανότητα να επιλέξουμε μόνο άνδρες;

Ορίζουμε τον Δ.Χ. Ω ως το σύνολο όλων των δυνατών (μη-διατεταγμένων) επιλογών 5 ατόμων από 100 (εφόσον σ΄ αυτό το πρόβλημα δεν μας απασχολεί η σειρά με την οποία επιλέγονται). Όπως είδαμε, $|\Omega| = \binom{100}{5} = 75287520$.

Ορίζουμε επίσης και τα τρία ενδεχόμενα

$$\begin{aligned} A &= \{\text{όλες οι 5άδες που αποτελούνται από 2 άνδρες και 3 γυναίκες}\}, \\ B &= \{\text{όλες οι 5άδες που αποτελούνται από 4 άνδρες και 1 γυναίκα}\}, \\ C &= \{\text{όλες οι 5άδες που αποτελούνται από μόνο άνδρες}\}. \end{aligned}$$

Θα υπολογίσουμε το πλήθος των στοιχείων τους.

Για το A, παρατηρούμε ότι το πείραμα μπορεί να χωριστεί σε δύο μέρη. Από το Λήμμα 2.3 μπορούμε να επιλέξουμε 2 άνδρες απ΄ τους 40 με $\binom{40}{2}$ τρόπους, και 3 γυναίκες απ΄ τις 60 με $\binom{60}{3}$ τρόπους. Άρα, ο συνδυασμός αυτών των δύο επιλογών, βάσει της πολλαπλασιαστικής αρχής, έχει

$$|A| = \binom{40}{2}\binom{60}{3} = \frac{40!60!}{2!38!3!57!} = \frac{40 \times 39 \times 60 \times 59 \times 58}{2 \times 3 \times 2} = 26691600$$

δυνατά αποτελέσματα. Με την ίδια συλλογιστική έχουμε, από το Λήμμα 2.3 και την πολλαπλασιαστική αρχή, ότι

$$|B| = \binom{40}{4}\binom{60}{1} = \cdots = 5483400, \qquad |C| = \binom{40}{5} = \cdots = 658008.$$

(Παρατηρήστε πως στον υπολογισμό του $|C|$ δεν πολλαπλασιάσαμε με τους τρόπους που μπορούμε να πάρουμε 0 γυναίκες από τις 60, αλλά αυτό δεν επηρεάζει το αποτέλεσμα διότι το πλήθος των τρόπων αυτών είναι $\binom{60}{0} = \frac{60!}{0!60!} = 1$.)

Αφού η επιλογή θεωρούμε ότι γίνεται «εντελώς τυχαία», μπορούμε να εφαρμόσουμε το Λήμμα 1.2 έτσι ώστε

$$\begin{aligned} P(A) &= \frac{|A|}{|\Omega|} = \frac{26691600}{75287520} \simeq 0.354, \\ P(B) &= \frac{|B|}{|\Omega|} = \frac{5483400}{75287520} \simeq 0.0728, \\ P(C) &= \frac{|C|}{|\Omega|} = \frac{658008}{75287520} \simeq 0.0087. \end{aligned}$$

Παράδειγμα 2.8. Σε κάποιες εκλογές είναι υποψήφιοι 3 φοιτητές και 7 καθηγητές. Εκλέγονται τυχαία 3 άτομα και ζητάμε την πιθανότητα να εκλεγούν τουλάχιστον ένας φοιτητής και τουλάχιστον ένας καθηγητής. Έστω X το πλήθος των φοιτητών που επιλέχθηκαν, και Y το πλήθος των καθηγητών. Πάλι με την ίδια συλλογιστική όπως στα δύο πιο πάνω παραδείγματα, εφόσον δεν μας ενδιαφέρει η διάταξη των τριών ατόμων, υπολογίζουμε την ζητούμενη πιθανότητα ως εξής:

$$\begin{aligned}
&P(X \geq 1,\ Y \geq 1) \\
&\quad = P\left(\{X=1,\ Y=2\} \cup \{X=2,\ Y=1\}\right) \\
&\quad = P\left(\{X=1,\ Y=2\}\right) + P\left(\{X=2,\ Y=1\}\right) \\
&\quad = \frac{\binom{3}{1}\binom{7}{2}}{\binom{10}{3}} + \frac{\binom{3}{2}\binom{7}{1}}{\binom{10}{3}} = \frac{21}{40} + \frac{7}{40} = 0.7.
\end{aligned}$$

Στην δεύτερη ισότητα χρησιμοποιήσαμε το ότι τα ενδεχόμενα είναι ξένα μεταξύ τους.

Ας υποθέσουμε τώρα πως, ανάλογα με τη σειρά εκλογής, αυτοί που εκλέγονται παίρνουν διαφορετικούς ρόλους σε μια επιτροπή – ο πρώτος γίνεται πρόεδρος, ο δεύτερος γραμματέας και ο τρίτος ταμίας. Ποια είναι η πιθανότητα του ενδεχόμενου A να εκλεγεί φοιτητής πρόεδρος, και καθηγητές γραμματέας και ταμίας; Εφόσον εδώ μας απασχολεί και η σειρά με την οποία επιλέγονται τα «αντικείμενα» (δηλαδή τα μέλη της επιτροπής), υπολογίζουμε αυτή την πιθανότητα βάσει του νέου δειγματικού χώρου Ω ο οποίος περιέχει όλες τις *διατεταγμένες* 3άδες, δηλαδή περιέχει

$$|\Omega| = 10!/(10-3)! = 720$$

στοιχεία. Το πλήθος των τριάδων που ανήκουν στο A είναι $3 \times 7 \times 6$, αφού έχουμε 3 επιλογές για τον φοιτητή που θα γίνει πρόεδρος, 7 επιλογές για τον καθηγητή που θα γίνει γραμματέας, και 6 επιλογές για τον καθηγητή που θα γίνει ταμίας (ανάμεσα στους 6 που μένουν).

Άρα, η πιθανότητα του ενδεχομένου που μας ενδιαφέρει ισούται με

$$P(A) = \frac{3 \times 7 \times 6}{720} = \frac{7}{40} \simeq 0.175.$$

2.3 Πολυωνυμικός Συντελεστής

Παράδειγμα 2.9. Πρέπει 12 διεργασίες να κατανεμηθούν σε 3 επεξεργαστές, δίνοντας τέσσερις διεργασίες στον κάθε επεξεργαστή. Με πόσους τρόπους μπορεί να γίνει αυτός ο καταμερισμός;

Για να απαντήσουμε, χωρίζουμε το πρόβλημα σε τρία μέρη. Αρχικά επιλέγουμε 4 διεργασίες απ΄ τις 12 για τον πρώτο επεξεργαστή, πράγμα που (βάσει του Λήμματος 2.3) μπορεί να γίνει με $\binom{12}{4}$ τρόπους. Κατόπιν, επιλέγουμε 4 διεργασίες απ΄ τις υπόλοιπες 8 για τον δεύτερο επεξεργαστή, πράγμα που μπορεί να γίνει με $\binom{8}{4}$ τρόπους. Και τέλος οι 4 διεργασίες που απομένουν πηγαίνουν στον τρίτο επεξεργαστή. Χρησιμοποιώντας την πολλαπλασιαστική αρχή, συνολικά αυτός ο καταμερισμός μπορεί να γίνει με

$$\binom{12}{4}\binom{8}{4} = \frac{12!}{4!8!}\frac{8!}{4!4!} = \frac{12!}{4!4!4!} \quad \text{τρόπους.}$$

Πιο γενικά, έχουμε το ακόλουθο λήμμα:

Λήμμα 2.4. (Πολυωνυμικός συντελεστής) *Για να μοιραστούν N αντικείμενα σε M ομάδες, όπου η πρώτη αποτελείται από k_1 αντικείμενα, η δεύτερη από k_2 αντικείμενα, κ.ο.κ., ως την ομάδα M η οποία αποτελείται από k_M αντικείμενα, υπάρχουν*

$$\binom{N}{k_1, k_2, \ldots, k_M} \triangleq \frac{N!}{k_1!k_2!\ldots k_M!}$$

δυνατοί συνδυασμοί. Η άνω ποσότητα καλείται πολυωνυμικός συντελεστής. (Φυσικά πρέπει να ισχύει ότι $k_1 + k_2 + \cdots + k_M = N$).

Απόδειξη. Κατά την επιλογή της πρώτης ομάδας, έχουμε N αντικείμενα και πρέπει να επιλέξουμε τα k_1. Αυτό μπορεί να γίνει με $N_1 = \binom{N}{k_1}$ τρόπους. Κατά την επιλογή της δεύτερης ομάδας, έχουμε $N - k_1$ αντικείμενα, από τα οποία πρέπει να επιλεγούν k_2. Αυτό μπορεί να γίνει με $N_2 = \binom{N-k_1}{k_2}$ τρόπους. Συνεχίζοντας με τον ίδιο τρόπο, προκύπτει ότι

$$N_1 = \binom{N}{k_1},\ N_2 = \binom{N-k_1}{k_2}, \ldots,\ N_{M-1} = \binom{N-k_1-k_2-\cdots-k_{M-2}}{k_{M-1}}.$$

Συνεπώς, ο πολυωνυμικός συντελεστής δίνεται από:

$$N_1N_2\ldots N_{M-1} = \binom{N}{k_1} \times \binom{N-k_1}{k_2} \times \cdots \times \binom{N-k_1-k_2-\cdots-k_{M-2}}{k_{M-1}} =$$

$$\frac{N!}{k_1!(N-k_1)!} \times \frac{(N-k_1)!}{k_2!(N-k_1-k_2)!} \cdots \frac{(N-k_1-k_2-\cdots-k_{M-2})!}{k_{M-1}!k_M!} = \frac{N!}{k_1!k_2!\ldots k_M!}.$$

Στην προτελευταία ισότητα χρησιμοποιήσαμε το ότι $k_1 + k_2 + \cdots + k_M = N$. □

Παρατήρηση: Αν έχουμε μόνο $M = 2$ ομάδες και N αντικείμενα, τότε για $k_1 = k$ αναγκαστικά θα έχουμε $k_2 = N - k$ και το Λήμμα 2.4 λέει πως υπάρχουν $\frac{N!}{k!(N-k)!}$ τρόποι να μοιράσουμε N αντικείμενα σε δύο ομάδες των k και $(N-k)$ αντίστοιχα. Αυτό είναι ταυτόσημο με το περιεχόμενο του Λήμματος 2.3, άρα το Λήμμα 2.4 αποτελεί γενίκευση του Λήμματος 2.3.

Παράδειγμα 2.10. Έχουμε 20 υπολογιστές, που αποτελούνται από 10 PC και 10 Apple, και τους μοιράζουμε τυχαία σε τρία clusters, που αποτελούνται από 10, 5 και 5 υπολογιστές, αντίστοιχα. Θα υπολογίσουμε τις πιθανότητες των ενδεχομένων

$$\begin{aligned} A &= \text{«όλα τα PC στο ίδιο cluster»}, \\ B &= \text{«4 PC στο πρώτο cluster, 3 PC στο δεύτερο και 3 PC στο τρίτο»}. \end{aligned}$$

Ο δειγματικός χώρος Ω που περιγράφει αυτό το πείραμα αποτελείται από όλους τους δυνατούς τρόπους με τους οποίους 20 αντικείμενα μπορούν να χωριστούν σε τρεις ομάδες των 10, 5 και 5 αντικειμένων. Άρα, από το Λήμμα 2.4 έχουμε

$$|\Omega| = \binom{20}{10,5,5} = \frac{20!}{10!5!5!} = 46558512 \simeq 46.5 \times 10^6.$$

Επιπλέον, οι υπολογιστές κατατάσσονται σε clusters τυχαία, οπότε υποθέτουμε ότι όλα τα αποτελέσματα είναι ισοπίθανα και θα χρησιμοποιήσουμε και πάλι το Λήμμα 1.2 για να υπολογίσουμε την πιθανότητα του A και του B.

Για το A παρατηρούμε πως «όλα τα PC στο ίδιο cluster» είναι ακριβώς ισοδύναμο με το «όλα τα PC στο ΠΡΩΤΟ cluster». Άρα, το πλήθος των στοιχείων του A ισούται με το πλήθος των τρόπων που μπορούν τα 10 Apple να μοιραστούν σε δύο clusters με 5 το καθένα, δηλαδή $\binom{10}{5}$. Συνεπώς,

$$P(A) = \binom{10}{5} \Big/ \binom{20}{10,5,5} = \frac{1}{184756} \simeq 5.41 \times 10^{-6}.$$

Τέλος, για το B, έχουμε όλους τους δυνατούς τρόπους με τους οποίους τα 10 PC μπορούν να μοιραστούν σε 3 clusters με αναλογία 4-3-3, σε συνδυασμό με όλους τους τρόπους με τους οποίους τα 10 Apple μπορούν να μοιραστούν σε 3 clusters με αναλογία 6-2-2. Άρα, έχουμε:

$$\begin{aligned} P(B) &= \frac{|A|}{|\Omega|} = \binom{10}{4,3,3}\binom{10}{6,2,2} \Big/ \binom{20}{10,5,5} \\ &= \frac{10!10!10!5!5!}{20!4!3!3!6!2!2!} = \frac{287}{2525} \simeq 0.1136. \end{aligned}$$

2.4 Επαναληπτικές Διατάξεις και Επαναληπτικοί Συνδυασμοί

Ορισμός 2.3. (Επαναληπτικές διατάξεις) *Έστω ένα σύνολο X που αποτελείται από πεπερασμένο πλήθος διακριτών στοιχείων $|X| = N$. Καλούμε επαναληπτική διάταξη κάθε διατεταγμένη k-άδα στοιχείων του X, όχι απαραίτητα διακριτών μεταξύ τους (και συνεπώς έχουμε επανάθεση στοιχείων κατά την δημιουργία της k-άδας).*

Παράδειγμα 2.11. Σε μια τάξη με 20 μαθητές, ο δάσκαλος, που έχει αμνησία, ρωτά κάθε πρωί έναν μαθητή να πει το μάθημα. Με πόσους διαφορετικούς τρόπους μπορούν να εξεταστούν οι μαθητές σε ένα διάστημα τριάντα σχολικών ημερών; Έχουμε 20 επιλογές για την πρώτη μέρα, 20 επιλογές για τη δεύτερη, κ.ο.κ., και τελικά, με εφαρμογή της πολλαπλασιαστικής αρχής, προκύπτει ότι συνολικά υπάρχουν 20^{30} ενδεχόμενα.

Για τη γενική περίπτωση, έχουμε το ακόλουθο λήμμα (η απόδειξη είναι προφανής).

Λήμμα 2.5. (Πλήθος επαναληπτικών διατάξεων) *Το πλήθος όλων των δυνατών επαναληπτικών διατάξεων k αντικειμένων που επιλέγονται από N αντικείμενα με επανάθεση ισούται με N^k.*

Είδαμε πως, αν επιλέξουμε k αντικείμενα από N *χωρίς* επανάθεση, υπάρχουν $\binom{N}{k}$ ομάδες k αντικειμένων που μπορούμε να επιλέξουμε. Τι θα γινόταν αν επιλέγαμε k αντικείμενα ανάμεσα σε n, επιτρέποντας την επανάθεση (και χωρίς να έχει σημασία η σειρά επιλογής);

Ορισμός 2.4. (Επαναληπτικοί συνδυασμοί) *Έστω ένα σύνολο X που αποτελείται από πεπερασμένο πλήθος διακριτών στοιχείων $|X| = N$. Καλούμε επαναληπτικό συνδυασμό κάθε μη διατεταγμένη k-άδα στοιχείων του X, όχι απαραίτητα διακριτών μεταξύ τους.*

Λήμμα 2.6. *Το πλήθος όλων των επαναληπτικών συνδυασμών k αντικειμένων που επιλέγονται από N αντικείμενα ισούται με:*

$$\binom{N+k-1}{k} = \binom{N+k-1}{N-1}.$$

Απόδειξη. Κάθε επιλογή μας μπορεί να περιγραφεί με μια ακολουθία από ακριβώς $N-1$ καθέτους (/) και ακριβώς k αστερίσκους (*) σε αυθαίρετη σειρά. Ο αριθμός των φορών που επιλέξαμε το i-οστό αντικείμενο ισούται με τον αριθμό από αστερίσκους ανάμεσα στην κάθετο $i-1$ και την κάθετο i. Ειδικά για το πρώτο αντικείμενο, μετράμε

τους αστερίσκους πριν την πρώτη κάθετο, και για το τελευταίο αντικείμενο μετράμε τους αστερίσκους μετά την τελευταία κάθετο. Όμως, ο αριθμός των διαφορετικών ακολουθιών αυτού του τύπου είναι ίσος με τον αριθμό των τρόπων που μπορούμε να τοποθετήσουμε k αστερίσκους σε $N+k-1$ θέσεις, ή, ισοδύναμα, $N-1$ καθέτους σε $N+k-1$ θέσεις. Άρα ο ζητούμενος αριθμός από συνδυασμούς αυτού του τύπου είναι ίσος με

$$\binom{N+k-1}{k} = \binom{N+k-1}{N-1}.$$

□

Παράδειγμα 2.12. *(Παγωτό)* Για να φτιάξουμε ένα παγωτό, επιλέγουμε 4 μπάλες από 7 διαθέσιμες γεύσεις. Μας νοιάζει η σειρά που θα βάλουμε τις μπάλες, και επιτρέπεται να βάλουμε πολλές μπάλες από κάτι, π.χ., μπορούμε να βάλουμε 2 μπάλες φράουλας, μια στην αρχή (στον πάτο του κυπέλλου) και μια στο τέλος (στην κορυφή του κυπέλλου).

Παρατηρήστε πως έχουμε επαναληπτικές διατάξεις 4 αντικειμένων από 7, άρα το πλήθος των διαφορετικών παγωτών που μπορεί να φτιαχτούν είναι $7^4 = 2401$.

Έστω τώρα πως ο υπάλληλος που φτιάχνει το παγωτό επιλέγει μια από τις 2401 επαναληπτικές διατάξεις στην τύχη, χωρίς να δείχνει προτίμηση σε κάποια από αυτές. Αν μια από τις γεύσεις είναι η σοκολάτα, ποια η πιθανότητα να καταλήξουμε με i μπάλες σοκολάτας, όπου $i = 0, 1, 2, 3, 4$; Έστω καταρχήν C το πλήθος των μπαλών σοκολάτας.

Για την περίπτωση $i = 0$, παρατηρούμε ότι οι επαναληπτικές διατάξεις που δεν έχουν σοκολάτα είναι οι επαναληπτικές διατάξεις 4 αντικειμένων από 6 (αφού η σοκολάτα έχει αποκλειστεί. Άρα

$$P(C=0) = \frac{6^4}{7^4} = \frac{1296}{2401}.$$

Για την περίπτωση $i = 1$, παρατηρούμε πως η μπάλα σοκολάτας μπορεί να βρίσκεται σε $\binom{4}{1} = 4$ θέσεις, και οι επαναληπτικές διατάξεις για τις υπόλοιπες θέσεις είναι 6^3. Άρα

$$P(C=1) = \frac{4 \times 6^3}{7^4} = \frac{864}{2401}.$$

Παρομοίως,

$$\begin{aligned} P(C=2) &= \frac{\binom{4}{2} \times 6^2}{7^4} = \frac{216}{2401}, \\ P(C=3) &= \frac{\binom{4}{3} \times 6^1}{7^4} = \frac{24}{2401}, \end{aligned}$$

ενώ υπάρχει μόνο μια επαναληπτική διάταξη με 4 μπάλες σοκολάτας, οπότε

$$P(C=4) = \frac{1}{210}.$$

Παρατηρήστε πως

$$P(C=0)+P(C=1)+P(C=2)+P(C=3)+P(C=4)=1.$$

Παράδειγμα 2.13. *(Milkshake)* Για να φτιάξουμε ένα milkshake, επιλέγουμε 4 μπάλες παγωτού από 7 διαθέσιμες γεύσεις, χωρίς όμως να μας νοιάζει η σειρά που θα βάλουμε τις μπάλες, και επιτρέπεται να βάλουμε πολλές μπάλες από κάτι, π.χ., μπορούμε να βάλουμε δύο μπάλες φράουλας. Οι δυνατοί συνδυασμοί είναι

$$\binom{4+7-1}{4}=\binom{10}{4}=210.$$

Με το συμβολισμό της απόδειξης του Λήμματος 2.6, η ακολουθία $//*/*//**/$ σημαίνει ότι επιλέξαμε 1 μπάλα από το τρίτο υλικό, μια από το τέταρτο, και δύο από το έκτο. Παρατηρήστε πως, σε σχέση με το προηγούμενο παράδειγμα, η διαφορά είναι ότι τώρα η σειρά με την οποία επιλέγει τις γεύσεις ο υπάλληλος δεν έχει σημασία, αφού τα υλικά θα περάσουν από το blender.

Έστω τώρα πως ο υπάλληλος που φτιάχνει το milkshake επιλέγει έναν από τους 210 επαναληπτικούς συνδυασμούς στην τύχη, χωρίς να δείχνει προτίμηση σε κάποιον από αυτούς. Αν μια από τις γεύσεις είναι η σοκολάτα, ποια η πιθανότητα να καταλήξουμε με i μπάλες σοκολάτας, όπου $i=0,1,2,3,4$; Έστω καταρχήν C το πλήθος των μπαλών σοκολάτας.

Για την περίπτωση $i=0$, παρατηρούμε ότι οι δυνατοί συνδυασμοί που δεν έχουν σοκολάτα είναι οι επαναληπτικοί συνδυασμοί 4 αντικειμένων από $7-1=6$ (αφού δεν υπολογίζεται η σοκολάτα). Άρα

$$P(C=0)=\frac{\binom{4+6-1}{4}}{\binom{4+7-1}{4}}=\frac{126}{210}=\frac{3}{5}.$$

Για την περίπτωση $i=1$, παρατηρούμε πως οι δυνατοί συνδυασμοί που έχουν μια μπάλα σοκολάτας είναι οι επαναληπτικοί συνδυασμοί $4-1$ αντικειμένων (αφού η μια μπάλα είναι η μπάλα σοκολάτας) από $7-1=6$ (αφού δεν υπολογίζεται η σοκολάτα). Άρα

$$P(C=1)=\frac{\binom{3+6-1}{3}}{\binom{4+7-1}{4}}=\frac{126}{210}=\frac{4}{15}.$$

Παρομοίως,

$$\begin{aligned}P(C=2)&=\frac{\binom{2+6-1}{2}}{\binom{4+7-1}{4}}=\frac{126}{210}=\frac{1}{10},\\ P(C=3)&=\frac{\binom{1+6-1}{1}}{\binom{4+7-1}{4}}=\frac{6}{210}=\frac{1}{35},\end{aligned}$$

ενώ υπάρχει ένας μόνο συνδυασμός με 4 μπάλες σοκολάτας, άρα

$$P(C = 4) = \frac{1}{210}.$$

Παρατηρήστε πως

$$P(C = 0) + P(C = 1) + P(C = 2) + P(C = 3) + P(C = 4) = 1.$$

Παρατηρήστε πως οι πιθανότητες $P(C = i)$ που βρήκαμε σε αυτό το παράδειγμα διαφέρουν από το προηγούμενο. Μπορείτε να εξηγήσετε τι συμβαίνει;

2.5 Παραδείγματα

Κανόνες Αρίθμησης

1. Πολλαπλασιαστική Αρχή: Αν εκτελεστούν διαδοχικά k πειράματα, καθένα με $N_1, N_2 \ldots, N_k$ δυνατά αποτελέσματα, τότε το νέο, σύνθετο πείραμα έχει $N_1 \times N_2 \times \cdots \times N_k$ δυνατά αποτελέσματα.

2. Υπάρχουν
$$(N)_k \triangleq \frac{N!}{(N-k)!}$$
δυνατές διατάξεις k αντικειμένων που επιλέγονται από N. Υπάρχουν $N!$ δυνατές μεταθέσεις.

3. Υπάρχουν
$$\binom{N}{k} \triangleq \frac{N!}{k!(N-k)!}$$
δυνατοί συνδυασμοί k αντικειμένων που επιλέγονται από N.

4. Υπάρχουν
$$\binom{N}{k_1, k_2, \ldots, k_M} \triangleq \frac{N!}{k_1! k_2! \ldots k_M!}$$
δυνατοί συνδυασμοί βάσει των οποίων μπορούν να μοιραστούν N αντικείμενα σε M ομάδες, όπου η πρώτη αποτελείται από k_1 αντικείμενα, η δεύτερη από k_2 αντικείμενα, κλπ., ως την ομάδα M η οποία αποτελείται από k_M αντικείμενα. (Πρέπει $k_1 + k_2 + \cdots + k_M = N$.)

5. Υπάρχουν N^k επαναληπτικές (δηλαδή με επανάθεση) διατάξεις μήκους k (δηλαδή διατεταγμένες k-άδες) από N αντικείμενα.

6. Υπάρχουν
$$\binom{N+k-1}{k} = \binom{N+k-1}{N-1}.$$
επαναληπτικοί συνδυασμοί (δηλαδή με επανάθεση) k αντικειμένων από N.

Παράδειγμα 2.14. *(Μέτρημα)* Πόσες δυνατές εκδοχές υπάρχουν στο καθένα από τα παρακάτω πειράματα;

1. Εγκαθιστούμε ένα πρόγραμμα διαδοχικά σε 10 υπολογιστές.
2. Επιλέγουμε με τη σειρά 3 από 12 αντικείμενα, χωρίς επανατοποθέτηση.
3. Στρίβουμε ένα κέρμα 6 φορές.
4. Επιλέγουμε 20 άτομα από 100 για μια δημοσκόπηση, χωρίς επανατοποθέτηση.
5. Ρίχνουμε ένα ζάρι 7 φορές.
6. Βάζουμε 13 ανθρώπους να κάτσουν σε μια σειρά.
7. Μοιράζουμε με τη σειρά 8 φύλλα από μια συνηθισμένη τράπουλα 52 φύλλων.
8. Ρίχνουμε ένα κέρμα 4 φορές και ένα ζάρι 2 φορές.

Σε κάθε περίπτωση πρέπει να χρησιμοποιήσουμε τον σωστό κανόνα αρίθμησης. Συγκεκριμένα:

1. Έχουμε 10 επιλογές για τον πρώτο υπολογιστή, 9 για τον δεύτερο, κ.ο.κ. Το πλήθος των μεταθέσεων είναι $10! = 3628800$.
2. Το πλήθος των διατεταγμένων 3άδων είναι
$$\frac{12!}{(12-3)!} = \frac{12 \times 11 \times 10 \times 9!}{9!} = 1320.$$
3. Κάθε μία από τις 6 ρίψεις μπορεί να καταλήξει σε 2 αποτελέσματα, άρα το πλήθος των επαναληπτικών διατάξεων είναι $2 \times 2 \times 2 \times 2 \times 2 \times 2 = 2^6 = 32$.
4. Εφόσον δεν μας ενδιαφέρει η σειρά επιλογής, το πλήθος των συνδυασμών είναι
$$\binom{100}{20} = \frac{100!}{(100-20)!20!} \simeq 5.359 \times 10^{20}.$$
5. Κάθε ζαριά έχει 6 δυνατά αποτελέσματα, συνεπώς από την πολλαπλασιαστική αρχή έχουμε $6 \times 6 \times 6 \times 6 \times 6 \times 6 \times 6 = 6^7$ δυνατές επαναληπτικές διατάξεις.
6. Έχουμε 13 επιλογές για τον πρώτο που θα επιλέξουμε, 12 επιλογές για τον δεύτερο, κ.ο.κ. Συνεπώς έχουμε συνολικά $13 \times 12 \times 11 \times \cdots \times 1 = 13! = 6,227,020,800$ μεταθέσεις.

7. Παρατηρήστε ότι τα φύλλα μοιράζονται με τη σειρά, συνεπώς ψάχνουμε για διατεταγμένες 8άδες. Έχουμε 52 επιλογές για το πρώτο χαρτί, 51 επιλογές για το δεύτερο χαρτί, κ.ο.κ., συνεπώς έχουμε $\frac{52!}{(52-8)!} = 30,342,338,208,000$ διαφορετικές 8άδες.

8. Κάθε κέρμα έχει 2 δυνατά αποτελέσματα και κάθε ζάρι 6 δυνατά αποτελέσματα, άρα από την πολλαπλασιαστική αρχή έχουμε $2 \times 2 \times 2 \times 2 \times 6 \times 6 = 576$ δυνατές εκδοχές.

Παράδειγμα 2.15. *(Αναγραμματισμοί)* Πόσοι αναγραμματισμοί υπάρχουν για την λέξη «ΘΕΜΑ»; Για τη λέξη «ΘΕΜΑΤΑ»; Για τη λέξη «ΠΑΝΑΘΗΝΑΪΚΟΣ»;

Έχουμε, κατά περίπτωση:

1. Για την λέξη «ΘΕΜΑ» υπάρχουν 4! αναγραμματισμοί αφού όλα τα γράμματα είναι διαφορετικά.

2. Για την λέξη «ΘΕΜΑΤΑ», όμως, υπάρχουν $6!/2! = 360$ αναγραμματισμοί. Πράγματι, 6! είναι όλες οι μεταθέσεις των 6 γραμμάτων. Όμως για το «Α» δεν μας νοιάζει η διάταξη. Οι διατάξεις των «Α» είναι 2!=2. Άρα πρέπει να διαιρέσουμε με 2!

3. Για τη λέξη «ΠΑΝΑΘΗΝΑΪΚΟΣ», παρατηρούμε πως το Α εμφανίζεται τρεις φορές, ενώ το Ν δύο. Όλα τα άλλα γράμματα εμφανίζονται μια φορά. Θα υπήρχαν 12! μεταθέσεις αν τα Α και Ν ήταν διακριτά. Αφού δεν είναι, πρέπει να διαιρέσουμε με $3! = 6$ και $2! = 2$ αντίστοιχα, και τελικά προκύπτει πως έχουμε $12!/(3!2!) = 39916800$ αναγραμματισμούς.

Παράδειγμα 2.16. *(Παράγκα)* Η ομάδα μας παίζει στην έδρα της 9 διαδοχικά παιχνίδια με αντίπαλες ομάδες. Έχει προαποφασιστεί να έχουμε 4 νίκες, 3 ισοπαλίες, και 2 ήττες. Με πόσους τρόπους μπορεί να γίνει αυτό;

Εδώ απλά θέλουμε να μοιράσουμε 9 «αντικείμενα» (τα 9 παιχνίδια) σε τρεις κατηγορίες των τεσσάρων, τριών και δύο. Συνεπώς, υπάρχουν,

$$\binom{9}{4,3,2} = \frac{9!}{4!3!2!} = 1260$$

τρόποι. Εναλλακτικά, αν δεν θέλουμε να χρησιμοποιήσουμε έναν έτοιμο τύπο, σκεφτόμαστε ως εξής: η πρώτη νίκη μπορεί να εμφανιστεί σε οποιονδήποτε από 9 αγώνες, η δεύτερη σε οποιονδήποτε από 8, κ.ο.κ., και για τα 9 αποτελέσματα, λαμβάνοντας τελικά 9! δυνατά αποτελέσματα. Παρατηρήστε ότι οι νίκες εμφανίζονται με 4! διαφορετικές σειρές, οι ισοπαλίες με 3! διαφορετικές σειρές, και οι ήττες με 2! διαφορετικές σειρές. Για κάθε μια συγκεκριμένη ακολουθία αποτελεσμάτων όπου οι νίκες/ισοπαλίες/ήττες

δεν είναι διακριτές, υπάρχουν $4! \times 3! \times 2!$ ακολουθίες όπου οι νίκες/ισοπαλίες/ήττες είναι διακριτές. Άρα, αν X ο ζητούμενος αριθμός, θα πρέπει $X \times 4! \times 3! \times 2! = 9!$, άρα τελικά $X = \frac{9!}{4!3!2!} = 1260$.

Παράδειγμα 2.17. *(Φοιτητικό Δωμάτιο)* Για να στολίσουμε ένα φοιτητικό δωμάτιο, πρέπει να αγοράσουμε 15 αρκουδάκια επιλέγοντας από 5 διαφορετικά είδη, και μπορούμε να αγοράσουμε πολλές φορές το ίδιο είδος αρκουδάκι. Η σειρά δεν έχει σημασία. Οι δυνατοί συνδυασμοί είναι

$$\binom{15+5-1}{15} = \binom{19}{15} = 3876.$$

Με το συμβολισμό της απόδειξης του Λήμματος 2.6, η ακολουθία

$$*****/****/******//$$

σημαίνει ότι επιλέξαμε 5 αρκουδάκια του πρώτου είδους, 4 του δεύτερου, και 6 του τρίτου.

Παράδειγμα 2.18. *(Poker)* Ένας παίκτης του πόκερ παίρνει 5 φύλλα από μια κανονική τράπουλα 52 φύλλων. Ποια είναι η πιθανότητα να έχει:

1. Καρέ (δηλαδή 4 ίδια φύλλα, για παράδειγμα 4 άσσους, 4 ντάμες, κτλ.);

2. Φουλ (δηλαδή ένα ζευγάρι και μια τριάδα, για παράδειγμα 3 άσσους και 2 ρηγάδες);

3. Χρώμα (δηλαδή όλα κούπες ή όλα σπαθιά ή όλα μπαστούνια ή όλα καρώ);

Λύση: Σε αυτό το παράδειγμα, ο δειγματικός χώρος Ω αποτελείται από όλες τις $\binom{52}{5}$ δυνατές 5άδες φύλλων που μπορεί να έχει ο παίκτης, αφού η σειρά δεν μας ενδιαφέρει.

1. Έστω A το σύνολο των 5άδων που περιέχουν καρέ. Κατ' αρχάς παρατηρούμε πως υπάρχουν 48 διαφορετικές 5άδες με καρέ του άσσου, μία για κάθε ένα από τα 48 φύλλα με τα οποία συμπληρώνουμε την πεντάδα. Υπάρχουν άλλες 48 5άδες με καρέ του Ρήγα, κ.ο.κ. Και, εφόσον υπάρχουν 13 διαφορετικά καρέ, το συνολικό πλήθος των στοιχείων του A είναι 13×48. Άρα, η ζητούμενη πιθανότητα είναι

$$P(A) = \frac{|A|}{|\Omega|} = \frac{13 \times 48}{\binom{52}{5}} = \frac{13 \times 48 \times 5! \times 47!}{52!} = \frac{1}{4165} \simeq 2.4 \times 10^{-4}.$$

2. Έστω B το σύνολο των 5άδων που αντιστοιχούν σε φουλ. Κατ' αρχάς, πόσες 5άδες αντιστοιχούν σε φουλ του άσσου με ρηγάδες; Υπάρχουν $\binom{4}{3}$ τρόποι να επιλέξουμε 3 άσσους και $\binom{4}{2}$ τρόποι να επιλέξουμε 2 ρηγάδες. Άρα υπάρχουν $\binom{4}{3} \times \binom{4}{2}$ 5άδες που περιέχουν φουλ του άσσου με ρηγάδες. Επιπλέον, υπάρχουν

13×12 διαφορετικά «είδη» φουλ αφού έχουμε 13 επιλογές για το νούμερο (ή φιγούρα) που θα εμφανιστεί σε τριάδα, και 12 επιλογές για το νούμερο (ή φιγούρα) που θα εμφανιστεί σε δυάδα. Άρα, το πλήθος όλων των δυνατών 5άδων που αντιστοιχούν σε φουλ είναι $|B| = 13 \times 12 \times \binom{4}{3} \times \binom{4}{2}$, και η πιθανότητα για φουλ είναι

$$P(B) = \frac{|B|}{|\Omega|} = \frac{13 \times 12 \times \binom{4}{3} \times \binom{4}{2}}{\binom{52}{5}} = \frac{13 \times 12 \times 4! \times 4! \times 5! \times 47!}{3! \times 1! \times 2! \times 2! \times 52!} = \frac{6}{4165} \simeq 0.0014.$$

3. Τέλος, έστω C το σύνολο των 5άδων που αντιστοιχούν σε χρώμα. Υπάρχουν $\binom{13}{5}$ 5άδες από κούπες, και το ίδιο πλήθος για τα μπαστούνια, σπαθιά και καρώ, άρα, $|C| = 4 \times \binom{13}{5}$. Άρα, η πιθανότητα του χρώματος είναι

$$P(C) = \frac{|C|}{|\Omega|} = \frac{4 \times \binom{13}{5}}{\binom{52}{5}} = \frac{4 \times 13! \times 5! \times 47!}{5! \times 8! \times 52!} = \frac{33}{16660} \simeq 0.002.$$

Παράδειγμα 2.19. *(Χορευτική Ομάδα)* Ένας ιμπρεσάριος μιας ομάδας 6 χορευτριών ετοιμάζει ένα σόου τριών χορών, με κάθε χορό να χορεύεται από ένα ζεύγος χορευτριών. Πόσες παραστάσεις μπορεί να ετοιμάσει ο ιμπρεσάριος, σε κάθε μια από τις ακόλουθες περιπτώσεις;

1. Οι χοροί εκτελούνται διαδοχικά, και σε κάθε χορό οι χορεύτριες έχουν διαφορετικά κουστούμια.

2. Οι χοροί εκτελούνται διαδοχικά, και σε κάθε χορό οι χορεύτριες έχουν ίδια κουστούμια.

3. Οι χοροί είναι όμοιοι, εκτελούνται ταυτόχρονα, και όλες οι χορεύτριες έχουν τα ίδια κουστούμια.

Έχουμε, σε κάθε περίπτωση:

1. Έχουμε 6 επιλογές για το ποια θα φορέσει το πρώτο κουστούμι του πρώτου χορού, 5 επιλογές για το ποια θα φορέσει το δεύτερο κουστούμι του πρώτου χορού, 4 επιλογές για το πρώτο κουστούμι του δεύτερου χορού, κ.ο.κ. Τελικά έχουμε 6! διαφορετικές παραστάσεις.

2. Έχουμε $\binom{6}{4}$ επιλογές για το πρώτο ζευγάρι και $\binom{4}{2}$ επιλογές για το δεύτερο ζευγάρι, άρα τελικά $\binom{6}{4} \times \binom{4}{2} = \frac{6!}{2^3} = 90$ διαφορετικές παραστάσεις.

3. Μια λύση είναι η εξής: Έστω πως οι χορεύτριες επιλέγουν την παρτενέρ τους με μια καθορισμένη σειρά, π.χ. βάσει αρχαιότητας. Υπάρχουν 5 επιλογές για την επιλογή παρτενέρ της πρώτης χορεύτριας που θα διαλέξει, και 3 επιλογές για την επόμενη που θα διαλέξει. Σύνολο έχουμε 15 επιλογές. Εναλλακτικά, μπορούμε να σκεφτούμε ως εξής: από το προηγούμενα σκέλος προκύπτει πως έχουμε 90 συνδυασμούς, αν έχει σημασία η σειρά των τριών ζευγαριών. Αν όμως δεν έχει σημασία η σειρά, πρέπει να διαιρέσουμε τους συνδυασμούς με το 3!, και τελικά παίρνουμε $90/3! = 15$ διαφορετικές παραστάσεις.

Παράδειγμα 2.20. *(Πέντε επί πέντε)* Με πόσους τρόπους μπορούν 10 άτομα να παίξουν ποδόσφαιρο 5×5; Παρατηρήστε πως υπάρχουν $\binom{10}{5} = 252$ τρόποι με τους οποίους μπορεί να σχηματιστεί η πρώτη ομάδα. Όμως, για κάθε ένα συνδυασμό με τον οποίο 5 άτομα εμφανίζονται στην πρώτη ομάδα, υπάρχει και άλλος ένας συνδυασμός με τον οποίο οι ίδιοι ακριβώς 5 εμφανίζονται στην δεύτερη ομάδα. Άρα τελικά, οι τρόποι με τους οποίους τα 10 άτομα μοιράζονται σε δύο ομάδες είναι $252/2 = 126$. Εναλλακτικά, φανταστείτε ότι ο ιδιοκτήτης της μπάλας επιλέγει την ομάδα του. Έχει συνολικά $\binom{9}{4} = 126$ επιλογές, και κάθε μια αντιστοιχεί σε μια δυνατή διαμέριση των 10 ατόμων.

Παρατηρήστε πως αν διαχωρίζαμε με κάποιο τρόπο τις ομάδες, για παράδειγμα λέγαμε πως η μια ομάδα θα κάνει τη σέντρα, ή η μια ομάδα φορά κόκκινα, κ.ο.κ., τότε θα υπήρχαν $\binom{10}{5} = 252$ τρόποι με τους οποίους μπορούσαν τα 10 άτομα να παίξουν ποδόσφαιρο.

Παράδειγμα 2.21. *(Ξενοδοχείο Ακρόπολις)* Έξι φίλοι συμφωνούν να συναντηθούν στο ξενοδοχείο Ακρόπολις των Αθηνών. Συμβαίνει όμως να υπάρχουν 4 ξενοδοχεία με το ίδιο όνομα. Κάθε ένας από τους 6 φίλους διαλέγει στην τύχη να πάει σε ένα από αυτά. Θα απαντήσουμε τα ακόλουθα ερωτήματα:

1. Ποια είναι η πιθανότητα να συναντηθούν ανά ζεύγη (εννοείται σε τρία διαφορετικά ξενοδοχεία);

2. Ποια είναι η πιθανότητα να βρεθούν δύο μόνοι τους και άλλοι τέσσερις σε δύο ζεύγη;

Καταρχήν, έστω πως δίνουμε έναν αριθμό από το 1 ως το 4 στα διαφορετικά ξενοδοχεία. Το τυχαίο πείραμα εδώ συνίσταται από την τυχαία επιλογή του ξενοδοχείου μετάβασης για κάθε ένα από τους 6 φίλους. Το σύνολο των δυνατών αποτελεσμάτων θα αποτελείται από όλες τις διατεταγμένες 6-άδες $(a_1, a_2, a_3, a_4, a_5, a_6)$ ξενοδοχείων (όπου κάθε $a_i = 1, 2, 3$ ή 4). Π.χ. το στοιχείο $(2, 3, 2, 1, 4, 4)$ αντιστοιχεί στο ενδεχόμενο ο πρώτος και ο τρίτος να πήγαν στο 2ο ξενοδοχείο, ο δεύτερος στο 3ο, ο τέταρτος στο 1ο, και οι πέμπτος και έκτος στο 4ο. Ο δειγματικός χώρος Ω επομένως θα αποτελείται από όλες τις $4^6 = 4096$ τέτοιες 6άδες.

Σχετικά με τις ζητούμενες πιθανότητες, έχουμε:

1. Έστω A το ενδεχόμενο να καταλήξουν οι φίλοι σε τρία ζεύγη. Παρατηρήστε ότι, πριν σκεφτούμε σε ποια ξενοδοχεία θα πάνε, υπάρχουν $\binom{6}{2}\binom{4}{2}$ τρόποι να μοιραστούν 6 άνθρωποι σε 3 ζεύγη, αν μας ενδιαφέρει η σειρά σχηματισμού. Συνεπώς υπάρχουν $\binom{6}{2}\binom{4}{2}/3!$ αν, όπως στην προκειμένη περίπτωση, δεν μας ενδιαφέρει η σειρά σχηματισμού.

   Αφού έχουμε τα τρία ζευγάρια, υπάρχουν $4 \times 3 \times 2 = 24$ τρόποι να μοιραστούν σε ξενοδοχεία (4 επιλογές για το πρώτο, 3 για το δεύτερο, 2 για το τρίτο). Άρα, το πλήθος των στοιχείων του A είναι, $|A| = 24 \times \binom{4}{2}\binom{6}{2}/3!$, και η ζητούμενη πιθανότητα ισούται με,

$$P(A) = \frac{|A|}{|\Omega|} = \frac{24 \times \binom{4}{2}\binom{6}{2}/3!}{4^6} = \frac{45}{512} \simeq 0.0879.$$

2. Έστω B το ενδεχόμενο να βρεθούν δύο μόνοι τους και άλλοι τέσσερις σε δύο ζεύγη. Παρατηρήστε ότι υπάρχουν $\binom{6}{2}\binom{4}{2}$ τρόποι να σχηματιστούν τα δύο ζεύγη, αν μας ενδιαφέρει η σειρά σχηματισμού, και $\binom{6}{2}\binom{4}{2}/2$ αν δεν μας ενδιαφέρει και ακολούθως $4 \times 3 \times 2 = 24$ τρόποι να μοιραστούν οι τέσσερις ομάδες (δύο ζεύγη και δύο άτομα μόνα) σε ξενοδοχεία (4 επιλογές για το πρώτο, 3 για το δεύτερο, 2 για το τρίτο). Άρα, η ζητούμενη πιθανότητα ισούται με:

$$P(B) = \frac{24 \times \binom{4}{2}\binom{6}{2}/2}{4^6} = \frac{135}{512} \simeq 0.2637.$$

Κεφάλαιο 3

Δεσμευμένη Πιθανότητα και Ανεξαρτησία

Ας πούμε πως ένας μετεωρολόγος μας πληροφορεί ότι, με βάση τα ιστορικά στατιστικά στοιχεία του καιρού στην Αθήνα, βρέχει μία στις 9 μέρες. Αν για κάποιο λόγο μας ενδιαφέρει τι καιρό κάνει τις Κυριακές (γιατί, π.χ., τις Κυριακές κάνουμε πικ-νικ στην Πάρνηθα), λογικά θα υποθέσουμε ότι μια στις 9 Κυριακές βρέχει. Αυτός ο συλλογισμός ισχύει γιατί έχουμε, έμμεσα, υποθέσει πως το αν βρέχει ή όχι σήμερα είναι *ανεξάρτητο* απ΄ το ποια μέρα της εβδομάδας είναι.

Αντίθετα, αν μας ενδιαφέρει ακριβώς αν θα βρέξει τις μέρες που έχει συννεφιά, θα ήταν λάθος να υποθέσουμε ότι μια στις 9 συννεφιασμένες μέρες βρέχει – το ποσοστό θα είναι προφανώς μεγαλύτερο, γιατί η συννεφιά *δεν είναι ανεξάρτητη* απ΄ την βροχή. Άρα, αν γνωρίζουμε ότι, π.χ., αύριο θα έχει συννεφιά, η πιθανότητα ότι θα βρέξει δεν θα είναι πλέον 1/9.

Το κεντρικό αντικείμενο αυτού του κεφαλαίου είναι η μαθηματική περιγραφή του πότε δύο ενδεχόμενα είναι *ανεξάρτητα*, και η χρήση της ανεξαρτησίας ενδεχόμενων στον υπολογισμό πιθανοτήτων.

3.1 Δεσμευμένη Πιθανότητα

Ορισμός 3.1. (Δεσμευμένη πιθανότητα) *Έστω δειγματικός χώρος Ω με ενδεχόμενα A, B, όπου $P(B) > 0$. Ορίζουμε την* <u>*δεσμευμένη πιθανότητα του A δεδομένου του B*</u> *ως*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Παρατήρηση: Η δεσμευμένη πιθανότητα του A δεδομένου του B εκφράζει την a posteriori πιθανότητα να συμβεί το A με δεδομένο ότι συνέβη το B. Η a priori πιθανότητα

είναι η $P(A)$. Ο ορισμός μπορεί να γίνει κατανοητός με ένα παράδειγμα.

Παράδειγμα 3.1. *(Εταιρία Πληροφορικής)* Μια εταιρία πληροφορικής απασχολεί 40 Έλληνες και 30 αλλοδαπούς εργαζόμενους, εκ των οποίων κάποιοι είναι τεχνικοί και κάποιοι προγραμματιστές:

	Έλληνες	αλλοδαποί
τεχνικοί	22	25
προγραμματιστές	18	5

Επιλέγουμε τυχαία έναν από τους 70 εργαζομένους, και εξετάζουμε τα ενδεχόμενα:

$$\begin{aligned} T &= \text{«επελέγη τεχνικός»}, \\ E &= \text{«επελέγη Έλληνας»}. \end{aligned}$$

Εφόσον η επιλογή είναι τυχαία, έχουμε

$$P(T) = \frac{22+25}{70} = \frac{47}{70} \quad \text{και} \quad P(E) = \frac{40}{70} = \frac{4}{7}.$$

Παρομοίως, η πιθανότητα να επιλέξουμε έναν Έλληνα τεχνικό είναι

$$P(E \cap T) = \frac{22}{70}.$$

Έστω τώρα πως γνωρίζουμε πως το άτομο που επελέγη είναι τεχνικός. Ποια η πιθανότητα να είναι Έλληνας; Από τον ορισμό της δεσμευμένης πιθανότητας, αυτό ισούται με

$$P(E|T) = \frac{P(E \cap T)}{P(T)} = \frac{22/70}{47/70} = \frac{22}{47}.$$

Σε αυτό το παράδειγμα μπορούμε να ελέγξουμε αν το αποτέλεσμα συμφωνεί με τη διαίσθησή μας για το τι θα πει «πιθανότητα ενός ενδεχομένου δεδομένου ενός άλλου» ως εξής: εφόσον γνωρίζουμε πως το επιλεγμένο άτομο είναι τεχνικός, διαισθητικά το πείραμά μας είναι ισοδύναμο με μια τυχαία επιλογή μεταξύ των $22 + 25 = 47$ τεχνικών. Και μια και απ᾿ αυτούς τους 47 οι 22 είναι Έλληνες, λογικά θα περιμέναμε η πιθανότητα του να επιλέξουμε έναν Έλληνα να ισούται με 22/47, το οποίο επιβεβαιώνεται απ᾿ τον προηγούμενό μας υπολογισμό που έγινε βάσει του ορισμού.

Τέλος, μπορούμε να εξετάσουμε την «αντίστροφη» περίπτωση: Ποια είναι η πιθανότητα να επιλέξαμε τεχνικό δεδομένου ότι επιλέξαμε κάποιον Έλληνα; Όπως και πριν, από τον ορισμό της δεσμευμένης πιθανότητας βρίσκουμε

$$P(T|E) = \frac{P(E \cap T)}{P(E)} = \frac{22/70}{40/70} = \frac{22}{40} = \frac{11}{20}.$$

Μπορείτε να επαληθεύσετε ότι και αυτό το αποτέλεσμα συμφωνεί με τη διαίσθησή σας;

Παράδειγμα 3.2. Μια οικογένεια με δύο παιδιά έχει τουλάχιστον ένα κορίτσι. Ποιο είναι το ενδεχόμενο να είναι και τα δύο κορίτσια; Για να απαντήσουμε το ερώτημα, επιλέγουμε σαν δειγματικό χώρο τον $\Omega = \{GG, GB, BG, BB\}$, όπου το αποτέλεσμα GB σημαίνει το πρώτο παιδί είναι κορίτσι (girl) και το δεύτερο αγόρι (boy), κ.ο.κ. Επίσης υποθέτουμε ότι όλα τα αποτελέσματα έχουν την ίδια πιθανότητα, 1/4. Άρα:

$$P(\{GG\}|\{BG, GB, GG\}) = \frac{P(\{GG\} \cap \{BG, GB, GG\})}{P(\{BG, GB, GG\})} = \frac{P(GG)}{P(\{BG, GB, GG\})} = \frac{1/4}{3/4} = 1/3.$$

Για να πειστείτε για την ορθότητα του αποτελέσματος, σκεφτείτε ως εξής: έστω μια πόλη με 1000 οικογένειες, εκ των οποίων 250 έχουν δύο αγόρια, 500 έχουν ένα αγόρι και ένα κορίτσι, και 250 έχουν δύο κορίτσια. Αν αποκλείσουμε τις 250 με τα δύο αγόρια και επιλέξουμε μια από τις άλλες στην τύχη, η πιθανότητα να επιλέξουμε μια οικογένεια με δύο κορίτσια είναι $250/750 = 1/3$.

Παράδειγμα 3.3. Επιλέγουμε 5 χαρτιά στην τύχη από μια συνηθισμένη τράπουλα. Ποια είναι η πιθανότητα του ενδεχόμενου A να επιλέξουμε τουλάχιστον μία φιγούρα;

$$P(A) = 1 - P(A') = 1 - \binom{40}{5} \Big/ \binom{52}{5} = 1 - \frac{40!47!}{35!52!} \simeq 0.746$$

Ποια είναι η πιθανότητα του ενδεχόμενου B να επιλέξουμε ακριβώς μια φιγούρα, δεδομένου ότι επιλέξαμε τουλάχιστον μία; Παρατηρήστε ότι $B \subset A$, επομένως

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)}{P(A)} = \binom{40}{4}\binom{12}{1} \Big/ \left[\binom{52}{5} P(A)\right] \simeq 0.565.$$

Παρατηρήσεις

1. Όπως δείχνουν τα άνω παραδείγματα, ουσιαστικά όταν υπολογίζουμε μια δεσμευμένη πιθανότητα $P(A|B)$, περιοριζόμαστε στο μικρότερο δειγματικό χώρο B.

2. Αντίθετα με τα άνω παραδείγματα, πολλές φορές μας δίνονται, ή είναι εύκολο να υπολογιστούν, κάποιες δεσμευμένες πιθανότητες, τις οποίες μπορούμε να χρησιμοποιήσουμε για να βρούμε κάποιες όχι δεσμευμένες πιθανότητες. Κλειδί για αυτές τις περιπτώσεις είναι το ακόλουθο λήμμα:

Λήμμα 3.1. (Πολλαπλασιαστικός τύπος) *Έστω δειγματικός χώρος Ω και μια ακολουθία ενδεχόμενων $A_1, A_2, \ldots, A_n$, όπου $P(A_1 A_2 \ldots A_{n-1}) > 0$. Ισχύει η ακόλουθη ισότητα:*

$$P(A_1 A_2 \ldots A_n) = P(A_1) P(A_2|A_1) P(A_3|A_2 A_1) \ldots P(A_n|A_1 A_2 \ldots A_{n-1}).$$

Απόδειξη. Καταρχήν, παρατηρούμε ότι αν $P(A_1A_2\dots A_{n-1}) > 0$, τότε θα έχουμε και $P(A_1A_2\dots A_i) > 0$ για κάθε $i = 1, 2, \dots, n-2$ (γιατί;) και έτσι όλες οι δεσμευμένες πιθανότητες είναι καλώς ορισμένες. Για να αποδείξουμε το ζητούμενο απλώς παρατηρούμε πως

$$\begin{aligned} & P(A_1)P(A_2|A_1)P(A_3|A_2A_1)\dots P(A_n|A_1A_2\dots A_{n-1}) \\ = \; & P(A_1) \times \frac{P(A_2A_1)}{P(A_1)} \times \frac{P(A_3A_2A_1)}{P(A_2A_1)} \times \dots \times \frac{P(A_1A_2\dots A_n)}{P(A_1A_2\dots A_{n-1})} \\ = \; & P(A_1A_2\dots A_n). \end{aligned}$$

Η τελευταία ισότητα προκύπτει με διαδοχικές απλοποιήσεις. □

Παράδειγμα 3.4. Αφαιρούμε διαδοχικά δύο μπάλες από ένα δοχείο με 10 άσπρες και 8 μαύρες. Ποια είναι η πιθανότητα να τραβήξουμε δύο άσπρες; Για να υπολογίσουμε την απάντηση, μπορούμε να χρησιμοποιήσουμε συνδυαστική: Υπάρχουν συνολικά $\binom{18}{2}$ συνδυασμοί, εκ των οποίων οι $\binom{10}{2}$ είναι οι συνδυασμοί των άσπρων μπαλών. Αν A_1 είναι το ενδεχόμενο να είναι η πρώτη μπάλα άσπρη, A_2 το ενδεχόμενο να είναι η δεύτερη μπάλα άσπρη, και συνεπώς A_1A_2 το ενδεχόμενο να βγάλουμε δύο άσπρες, η πιθανότητα του A_1A_2 είναι

$$P(A_1A_2) = \frac{\binom{10}{2}}{\binom{18}{2}} = \frac{10!16!2!}{18!8!2!} = \frac{10 \times 9}{18 \times 17} = \frac{5}{17}.$$

Πιο απλά όμως, παρατηρούμε πως

$$P(A_1A_2) = P(A_1)P(A_2|A_1) = \frac{10}{18} \times \frac{9}{17} = \frac{5}{17}.$$

Η δεσμευμένη πιθανότητα $P(A_2|A_1)$ ισούται με $\frac{9}{17}$ γιατί, με δεδομένο ότι έχει φύγει μια άσπρη μπάλα στην αρχή, ο νέος δειγματικός χώρος, A_1, αποτελείται από 17 αποτελέσματα, εκ των οποίων 9 ανήκουν στο $A_1 \cap A_2$. Παρατηρήστε πόσο πιο απλά μπορούμε να βρούμε το αποτέλεσμα, σε σχέση με την χρήση συνδυαστικής.

Ακολούθως, έστω M_1 το ενδεχόμενο να τραβήξουμε πρώτα μια μαύρη, και M_2 το ενδεχόμενο η δεύτερη να είναι μαύρη. Έστω το ενδεχόμενο να τραβήξουμε μια λευκή και μια μαύρη, δηλαδή το ενδεχόμενο $A_1M_2 \cup A_2M_1$. Η πιθανότητα αυτού του ενδεχόμενου είναι

$$\begin{aligned} P(A_1M_2 \cup M_1A_2) = P(A_1M_2) + P(M_1A_2) &= P(A_1)P(M_2|A_1) + P(M_1)P(A_2|M_1) \\ &= \frac{10}{18} \times \frac{8}{17} + \frac{8}{18} \times \frac{10}{17} = \frac{80}{153}. \end{aligned}$$

Στο ίδιο αποτέλεσμα μπορούμε να καταλήξουμε με χρήση συνδυαστικής, και πάλι όμως με περισσότερες πράξεις.

Παράδειγμα 3.5. *(Monty Hall — συνέχεια)* Σε συνέχεια του προβλήματος του Monty Hall του Παραδείγματος 1.11, τι είναι καλύτερο: να αλλάζει, ή αν μην αλλάζει κουρτίνα ο παρουσιαστής; Θα απαντήσουμε αυτό το ερώτημα υπολογίζοντας τις πιθανότητες όλων των αποτελεσμάτων του Δ.Χ. του Σχήματος 1.4, και στις δύο περιπτώσεις.

Υπενθυμίζουμε ότι τα αποτελέσματα είναι τριάδες της μορφής (X, X, X), όπου τα X παίρνουν τιμές A, B, ή C, και το πρώτο στοιχείο δείχνει την επιλογή του διαγωνιζόμενου, το δεύτερο την κουρτίνα που αποκαλύφθηκε, και το τρίτο την κουρτίνα που επέλεξε τελικά ο διαγωνιζόμενος. Το δώρο βρίσκεται στην κουρτίνα A.

Για να υπολογίσουμε τις πιθανότητες, χρειαζόμαστε δυο υποθέσεις. Πρώτον, υποθέτουμε πως η αρχική επιλογή του παίκτη είναι τυχαία, και μάλιστα πως τα ενδεχόμενα «ο παίκτης πρώτα επιλέγει την A»,«ο παίκτης πρώτα επιλέγει την B» και «ο παίκτης πρώτα επιλέγει την C» έχουν πιθανότητα $1/3$ το καθένα. Επιπλέον, υποθέτουμε ότι όταν ο παρουσιαστής έχει επιλογή στο ποια κουρτίνα θα ανοίξει, ανοίγει κάθε μια από τις δύο με πιθανότητα $1/2$.

Στρατηγική 1: Έστω πως ο παίκτης αλλάζει κουρτίνα με πιθανότητα 1. Τότε ο παίκτης θα κερδίσει αν (και μόνο αν) αρχικά επιλέξει την κουρτίνα B ή την κουρτίνα C. Στο Σχήμα 3.1 (άνω) φαίνεται ο δειγματικός χώρος και οι πιθανότητες όλων των αποτελεσμάτων. Τα αποτελέσματα που οδηγούν σε νίκη είναι υπογραμμισμένα, και έχουν ολική πιθανότητα $2/3$.

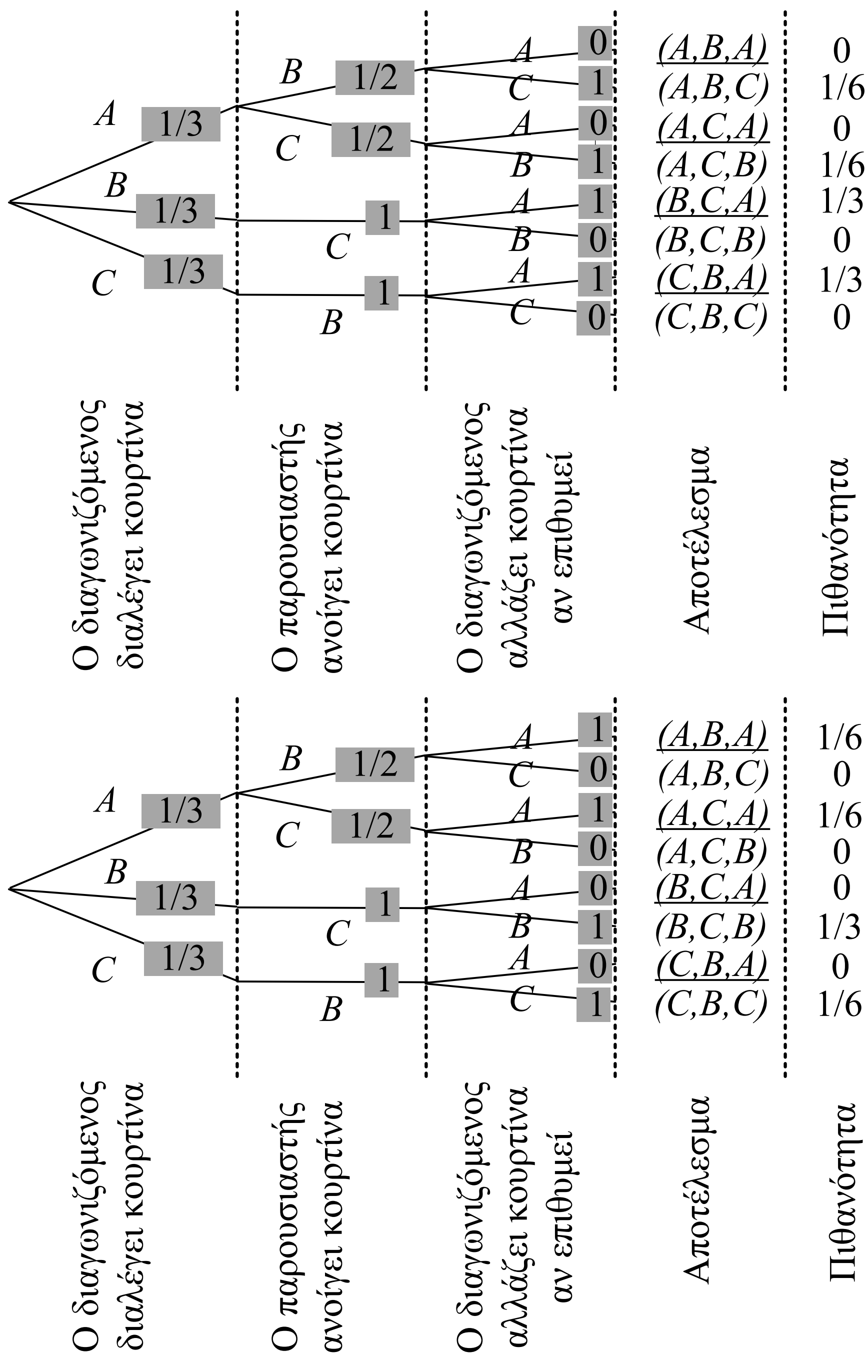
Στρατηγική 2: Ο παίκτης αλλάζει κουρτίνα με πιθανότητα 0. Τότε ο παίκτης θα κερδίσει αν (και μόνο αν) αρχικά επιλέξει την κουρτίνα A. Στο Σχήμα 3.1 (κάτω) φαίνεται ο δειγματικός χώρος και οι πιθανότητες όλων των αποτελεσμάτων. Τα αποτελέσματα που οδηγούν σε νίκη είναι υπογραμμισμένα, και έχουν ολική πιθανότητα $1/3$.

Συνεπώς, *συμφέρει τον παίκτη να αλλάζει κουρτίνα*!

Οι υπολογισμοί των πιθανοτήτων όλων των αποτελεσμάτων έγιναν με χρήση του πολλαπλασιαστικού τύπου. Για παράδειγμα, ας εξετάσουμε την πιθανότητα να προκύψει το αποτέλεσμα (A, B, C) αν ο παίκτης ακολουθεί την πρώτη στρατηγική. Έστω $(A, *, *)$ το ενδεχόμενο να επιλέξει ο παίκτης την κουρτίνα A, $(*, B, *)$ το ενδεχόμενο να ανοίξει ο παρουσιαστής την κουρτίνα B, και $(*, *, C)$ το ενδεχόμενο να καταλήξει ο παίκτης με την κουρτίνα C. (Από ποια αποτελέσματα αποτελείται το κάθε ενδεχόμενο;). Έχουμε:

$$
\begin{aligned}
&P((A, B, C)) \\
&\quad = P((A, *, *) \cap (*, B, *) \cap (*, *, C)) \\
&\quad = P((A, *, *)) \times P((*, B, *)|(A, *, *)) \times P((*, *, C)|(A, *, *) \cap (*, B, *)) \\
&\quad = \frac{1}{3} \times \frac{1}{2} \times 1 = \frac{1}{6}.
\end{aligned}
$$

Στην τρίτη ισότητα, χρησιμοποιήσαμε τις δύο υποθέσεις μας και το γεγονός ότι ο παίκτης ακολουθεί την πρώτη στρατηγική.



Σχήμα 3.1: Ο δειγματικός χώρος και οι πιθανότητες όλων των αποτελεσμάτων του Παραδείγματος 3.5 για τις δύο διαφορετικές στρατηγικές: 1) ο παίκτης αλλάζει πάντα κουρτίνα, και 2) ο παίκτης δεν αλλάζει ποτέ κουρτίνα. Παρατηρήστε πως η πιθανότητα ενός αποτελέσματος ισούται με το γινόμενο των πιθανοτήτων όλων των κλάδων του δέντρου που οδηγούν στο εν λόγω αποτέλεσμα.

Παράδειγμα 3.6. *(Poker με δεσμευμένες πιθανότητες)* Ένας παίκτης του πόκερ παίρνει 5 φύλλα από μια κανονική τράπουλα 52 φύλλων. Ποια είναι η πιθανότητα για:

1. Καρέ (δηλαδή 4 ίδια φύλλα, για παράδειγμα 4 άσσους, 4 ντάμες, κτλ.);
2. Φουλ (δηλαδή ένα ζευγάρι και μια τριάδα, για παράδειγμα 3 άσσους και 2 ρηγάδες);
3. Χρώμα (δηλαδή όλα κούπες ή όλα σπαθιά ή όλα μπαστούνια ή όλα καρώ);

Θα χρησιμοποιήσουμε την έννοια της δεσμευμένης πιθανότητας, και όχι κανόνες της συνδυαστικής όπως κάναμε στο Παράδειγμα 2.18.

1. Φανταστείτε πως μας αποκαλύπτονται τα φύλλα όχι ταυτόχρονα, όπως σε λύσεις που βασίζονται σε συνδυαστική, αλλά διαδοχικά. Παρατηρήστε ότι το ενδεχόμενο του καρέ μπορεί να γραφεί ως ένωση πέντε ξένων ενδεχόμενων E_i, όπου το E_i είναι το «παράταιρο» φύλλο να είναι το i-οστό που θα μας αποκαλυφθεί. Λόγω συμμετρίας, η ζητούμενη πιθανότητα είναι ίση με $5P(E_1)$ (γιατί;)

   Για να υπολογίσουμε το $P(E_1)$, σκεφτόμαστε ως εξής: Έστω πως έχουμε τραβήξει το πρώτο χαρτί. Η πιθανότητα να κάνουμε τελικά καρέ είναι η πιθανότητα η αξία (δηλαδή το «νούμερο», άσσος, 2, 3, κ.ο.κ.) του δεύτερου φύλλου να είναι διαφορετική από του πρώτου (48/51), επί την πιθανότητα το τρίτο να έχει την ίδια αξία με το δεύτερο (3/50), επί την πιθανότητα το τέταρτο να είναι ίδιο με το δεύτερο (2/49), επί την πιθανότητα το τελευταίο να είναι ίδιο με το δεύτερο (1/48).

   Αν θέλουμε να είμαστε πιο σχολαστικοί, ορίζουμε τα ακόλουθα ενδεχόμενα:

$$\begin{aligned} A_2 &= \text{Η αξία του δεύτερου φύλλου είναι διαφορετική απ' του πρώτου,} \\ A_3 &= \text{Το τρίτο φύλλο έχει ίδια αξία με το δεύτερο,} \\ A_4 &= \text{Το τέταρτο φύλλο έχει ίδια αξία με το δεύτερο,} \\ A_5 &= \text{Το πέμπτο φύλλο έχει ίδια αξία με το δεύτερο.} \end{aligned}$$

   Τότε, έχουμε, από τον ορισμό της δεσμευμένη πιθανότητας

$$\begin{aligned} P(E_1) &= P(A_2 \cap A_3 \cap A_4 \cap A_5) \\ &= P(A_2)P(A_3|A_2)P(A_4|A_2 \cap A_3)P(A_5|A_2 \cap A_3 \cap A_4) \\ &= \frac{48}{51} \times \frac{3}{50} \times \frac{2}{49} \times \frac{1}{48}. \end{aligned}$$

   Άρα, η πιθανότητα να έρθει καρέ είναι ίση με

$$5P(E_1) = 5 \times \frac{48}{51} \times \frac{3}{50} \times \frac{2}{49} \times \frac{1}{48} = \frac{1}{4165} \simeq 2.4 \times 10^{-4}.$$

2. Για να υπολογίσουμε την πιθανότητα φουλ, θα υπολογίσουμε πρώτα την πιθανότητα του ενδεχόμενου B_2 να σηκώσουμε φύλλα από ακριβώς δύο «αξίες» (δηλαδή νούμερα). Παρατηρήστε ότι έχουμε φύλλα με ακριβώς δύο διαφορετικές αξίες αν και μόνο αν έχουμε καρέ ή φουλ, οπότε το ενδεχόμενο B_2 μπορεί να εκφραστεί ως $B_2 = A \cup B$ όπου B=«έχουμε φουλ» και A=«έχουμε καρέ», με A, B ξένα. Την πιθανότητα του A την υπολογίσαμε ήδη, άρα για να βρούμε την ζητούμενη πιθανότητα $P(B) = P(B_2) - P(A)$, αρκεί να υπολογίσουμε $P(B_2)$.

   Για αυτό το σκοπό θα εκφράσουμε το B_2 ως την ένωση των ξένων ενδεχομένων
   $$B_2 = B_{22} \cup B_{23} \cup B_{24} \cup B_{25},$$
   όπου το B_{2i} είναι το υποσύνολο του B_2 όπου η δεύτερη απ΄ τις δύο αξίες της 5άδας εμφανίζεται για πρώτη φορά στο φύλλο i. Παρατηρήστε πως:
   $$P(B_{22}) = \frac{48}{51} \times \frac{6}{50} \times \frac{5}{49} \times \frac{4}{48}, \quad P(B_{23}) = \frac{3}{51} \times \frac{48}{50} \times \frac{5}{49} \times \frac{4}{48},$$
   $$P(B_{24}) = \frac{3}{51} \times \frac{2}{50} \times \frac{48}{49} \times \frac{4}{48}, \quad P(B_{25}) = \frac{3}{51} \times \frac{2}{50} \times \frac{1}{49} \times \frac{48}{48}.$$

   Άρα τελικά, η ζητούμενη πιθανότητα $P(B)$ ισούται με
   $$P(B) = P(B_{22}) + P(B_{23}) + P(B_{24}) + P(B_{25}) - P(A) = \frac{6}{4165} \simeq 0.0014.$$

3. Έστω ότι έχουμε τραβήξει το πρώτο χαρτί. Η πιθανότητα το δεύτερο να είναι του ίδιου χρώματος είναι $\frac{12}{51}$, γιατί έχουν μείνει 51 φύλλα προς επιλογή, εκ των οποίων 12 είναι του ιδίου χρώματος. Με δεδομένο ότι το δεύτερο φύλλο έχει το ίδιο χρώμα με το πρώτο, το τρίτο φύλλο θα είναι επίσης του ίδιου χρώματος με πιθανότητα $\frac{11}{50}$, κ.ο.κ.

   Πιο συστηματικά, ορίζοντας τα ενδεχόμενα
   $$Z_i = \text{«Το } i \text{ φύλλο έχει ίδιο χρώμα με το πρώτο»}, \quad i = 2, 3, 4, 5,$$
   έχουμε
   $$\begin{aligned} P(\text{«χρώμα»}) &= P(Z_2 \cap Z_3 \cap Z_4 \cap Z_5) \\ &= P(Z_2)P(Z_3|Z_2)P(Z_4|Z_2 \cap Z_3)P(Z_5|Z_2 \cap Z_3 \cap Z_4) \\ &= \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48} = \frac{33}{16660} \simeq 0.002. \end{aligned}$$

Παρατήρηση: Συγκρίνοντας αυτό το παράδειγμα με το Παράδειγμα 2.18, μπορείτε να διαπιστώσετε ότι όταν ένα πρόβλημα λύνεται με χρήση συνδυαστικής ή χρήση δεσμευμένης πιθανότητας, συχνά η μια μέθοδος είναι απλούστερη και/ή περιλαμβάνει λιγότερες πράξεις από την άλλη.

3.2 Ιδιότητες Δεσμευμένης Πιθανότητας

Η δεσμευμένη πιθανότητα $P(A|B)$ εκφράζει ουσιαστικά την πιθανότητα του A στον καινούργιο δειγματικό χώρο B. Άρα, αναμένουμε να είναι μέτρο πιθανότητας. Πράγματι, ισχύει το ακόλουθο:

Λήμμα 3.2. (Η δεσμευμένη πιθανότητα είναι πιθανότητα) *Έστω δειγματικός χώρος Ω με μέτρο πιθανότητας P, και έστω ενδεχόμενο B με $P(B) > 0$. Η συνάρτηση $Q(A)$ η οποία ορίζεται ως $Q(A) = P(A|B)$ για κάθε ενδεχόμενο A είναι ένα μέτρο πιθανότητας στον δειγματικό χώρο Ω, δηλαδή ικανοποιεί τις συνθήκες:*

1. *$P(A|B) \geq 0$ για κάθε ενδεχόμενο A.*

2. *$P(\Omega|B) = 1$.*

3. *Αν τα A_i, $i = 1, 2, \ldots$ είναι ξένα ενδεχόμενα, τότε*

$$P\left(\bigcup_{i=1}^{\infty} A_i \Big| B\right) = \sum_{i=1}^{\infty} P(A_i|B).$$

Απόδειξη. 1. Εξ ορισμού $P(A|B) = \frac{P(A\cap B)}{P(B)}$. Από το πρώτο αξίωμα των πιθανοτήτων έχουμε $P(A \cap B) \geq 0$, και εξ υποθέσεως $P(B) > 0$, συνεπώς $\frac{P(A\cap B)}{P(B)} \geq 0$.

2. Αφού $B \subseteq \Omega$, θα έχουμε $P(\Omega|B) = \dfrac{P(\Omega \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$.

3. Έστω A, C, ξένα ενδεχόμενα.

$$\begin{aligned} P(A \cup C|B) = \frac{P((A \cup C) \cap B)}{P(B)} &= \frac{P((A \cap B) \cup (C \cap B))}{P(B)} \\ &= \frac{P(A \cap B) + P(C \cap B)}{P(B)} = P(A|B) + P(C|B). \end{aligned}$$

Η δεύτερη ισότητα προέκυψε από την επιμεριστική ιδιότητα. Η τρίτη ισότητα προέκυψε από το ότι τα $A \cap B$ και $C \cap B$ είναι ξένα, ως υποσύνολα ξένων ενδεχόμενων. Η ιδιότητα μπορεί εύκολα να γενικευτεί για την περίπτωση αριθμήσιμα άπειρων ξένων ενδεχόμενων.

□

Λήμμα 3.3. (Ιδιότητες δεσμευμένης πιθανότητας) *Έστω Ω ένας δειγματικός χώρος και P ένα μέτρο πιθανότητας. Στα ακόλουθα, A, B, C είναι ενδεχόμενα του Ω. Έστω επίσης ενδεχόμενο D με $P(D) > 0$. Ισχύουν οι ακόλουθες ιδιότητες:*

1. $P(A'|D) = 1 - P(A|D)$.
2. $P(\emptyset|D) = 0$.
3. $P(A|D) \leq 1$.
4. $A \subseteq B \Rightarrow P(A|D) \leq P(B|D)$.
5. $P(A \cap B'|D) = P(A|D) - P(A \cap B|D)$.
6. $P(A \cup B|D) = P(A|D) + P(B|D) - P(A \cap B|D)$.
7. $P(A \cup B|D) \leq P(A|D) + P(B|D)$.
8. $P(A \cup B \cup C|D) = P(A|D) + P(A' \cap B|D) + P(A' \cap B' \cap C|D)$.
9. $$P(A \cup B \cup C|D) = P(A|D) + P(B|D) + P(C|D) - P(A \cap B|D) - P(A \cap C|D) - P(B \cap C|D) + P(A \cap B \cap C|D).$$

Απόδειξη. Οι άνω ιδιότητες έχουν αποδειχθεί στο Λήμμα 1.4 για κάθε μέτρο πιθανότητας, άρα θα ισχύουν και για μέτρα πιθανότητας που προέρχονται από δέσμευση. □

Παράδειγμα 3.7. *(Επιπλέον ιδιότητες δεσμευμένης πιθανότητας)* Έστω δύο οποιαδήποτε ενδεχόμενα A, B. Θα υπολογίσουμε την δεσμευμένη πιθανότητα $P(A|B)$ αν $A \cap B = \emptyset$, αν $A \subseteq B$, και αν $B \subseteq A$.

Αν $A \cap B = \emptyset$, τότε τα A και B είναι αδύνατον να συμβούν μαζί, οπότε, διαισθητικά, δεδομένου ότι έχει συμβεί το B είναι αδύνατον να συμβεί και το A. Πράγματι,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = 0.$$

Αν $A \subseteq B$, τότε:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}.$$

Αν $B \subseteq A$, τότε, δεδομένου ότι έχει συμβεί το B σίγουρα θα έχει συμβεί κάποιο από τα αποτελέσματα που ανήκουν και στο A. Πράγματι,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

Παράδειγμα 3.8. Θα δείξουμε ότι

$$P(A \cup B \cup C) = 1 - P(A'|B' \cap C')P(B'|C')P(C').$$

Πράγματι,

$$P(A \cup B \cup C) = 1 - P((A \cup B \cup C)') = 1 - P(A' \cap B' \cap C'),$$

και παρατηρώντας ότι

$$P(A' \cap B' \cap C') = P(A'|B' \cap C')P(B'|C')P(C'),$$

προκύπτει τελικά ότι

$$P(A \cup B \cup C) = 1 - P(A'|B' \cap C')P(B'|C')P(C').$$

3.3 Ο Κανόνας του Bayes

Σε πολλά επιστημονικά αλλά και καθημερινά προβλήματα, συχνά προκύπτει το εξής ερώτημα: Αν γνωρίζουμε την τιμή μιας δεσμευμένης πιθανότητας $P(A|B)$, πως μπορούμε να υπολογίσουμε την πιθανότητα $P(B|A)$ όπου οι ρόλοι των ενδεχομένων A και B έχουν αντιστραφεί; Ο κανόνας του Bayes αποτελεί το βασικό εργαλείο για την επίλυση τέτοιου είδους προβλημάτων. Μια από τις σημαντικότερες και πιο συνηθισμένες τέτοιες περιπτώσεις αφορούν προβλήματα που σχετίζονται με τεστ διαγνωστικών εξετάσεων. Εκεί, κατά κανόνα γνωρίζουμε την στατιστική συμπεριφορά του τεστ κάτω από ελεγχόμενες συνθήκες, π.χ., ξέρουμε την πιθανότητα ένα τεστ εγκυμοσύνης να έχει θετικό αποτέλεσμα *δεδομένου ότι η εξεταζόμενη είναι έγκυος*. Αλλά για εκείνη που κάνει το τεστ το σημαντικό ερώτημα είναι «δεδομένου ότι το τεστ είναι θετικό, πόσο πιθανό είναι να είμαι έγκυος»; Επιπλέον, ο κανόνας του Bayes αποτελεί την αφετηρία μιας πολύ σημαντικής επιστημονικής περιοχής, της *Στατιστικής κατά Bayes*, η οποία είναι εξαιρετικά ενεργή ερευνητικά και τα τελευταία 20 περίπου χρόνια έχει βρει εφαρμογές σε κεντρικά θέματα της σύγχρονης πληροφορικής.

Παράδειγμα 3.9. Έστω πως, σε έναν πληθυσμό 10 εκατομμυρίων ανθρώπων, 20,000 άτομα είναι φορείς του ιού HIV. Επιλέγεται ένα άτομο τυχαία και του γίνεται μια εξέταση για HIV, η οποία έχει ποσοστό σφάλματος 5%, δηλαδή,

$$\begin{aligned} P(\text{«αποτέλεσμα εξέτασης θετικό»} \mid \text{«όχι φορέας του HIV»}) &= 0.05, \\ P(\text{«αποτέλεσμα εξέτασης αρνητικό»} \mid \text{«φορέας του HIV»}) &= 0.05. \end{aligned}$$

Η σημαντική ερώτηση εδώ για τον εξεταζόμενο είναι: Αν το αποτέλεσμα της εξέτασης είναι θετικό (δηλαδή υποστηρίζει πως ο εξεταζόμενος είναι φορέας του ιού), ποια είναι η πιθανότητα ο εξεταζόμενος να είναι πράγματι φορέας; Ίσως φαίνεται εκ πρώτης όψεως «προφανές» πως η απάντηση είναι 95%, αλλά, όπως θα δούμε, η σωστή απάντηση είναι πολύ διαφορετική.

Για να προσεγγίσουμε το πρόβλημα συστηματικά, ορίζουμε τα ενδεχόμενα

$$\begin{aligned} A &= \text{«αποτέλεσμα εξέτασης θετικό»}, \\ H &= \text{«ο εξεταζόμενος είναι φορέας του HIV»}, \end{aligned}$$

και καταγράφουμε τα δεδομένα του προβλήματος:

$$P(H) = \frac{20000}{10000000} = \frac{2}{1000} = 0.002, \quad P(A|H') = P(A'|H) = 0.05.$$

Η ζητούμενη πιθανότητα είναι η $P(H|A)$.

Ξεκινάμε υπολογίζοντας την πιθανότητα του ενδεχομένου A, δηλαδή του να βγει θετικό το αποτέλεσμα της εξέτασης ενός τυχαία επιλεγμένου ατόμου. Εκφράζοντας το

A ως την ένωση δύο ξένων ενδεχομένων, $A = (A \cap H) \cup (A \cap H')$, έχουμε

$$P(A) = P(A \cap H) + P(A \cap H'),$$

και χρησιμοποιώντας τον πολλαπλασιαστικό κανόνα,

$$\begin{aligned} P(A) &= P(A|H)P(H) + P(A|H')P(H') \\ &= (1 - P(A'|H))P(H) + P(A|H')(1 - P(H)). \end{aligned}$$

Αντικαθιστώντας τις τιμές των πιθανοτήτων,

$$P(A) = (1 - 0.05) \times 0.002 + 0.05 \times (1 - 0.002) \simeq 0.0518.$$

Τώρα είμαστε σε θέση να απαντήσουμε το βασικό μας ερώτημα: δεδομένου ότι η εξέταση βγήκε θετική, ποια η πιθανότητα να είναι φορέας του HIV ο εξεταζόμενος; Με διπλή εφαρμογή του ορισμού της δεσμευμένης πιθανότητας βρίσκουμε:

$$\begin{aligned} P(H|A) &= \frac{P(A \cap H)}{P(A)} = \frac{P(A|H)P(H)}{P(A)} = \frac{[1 - P(A'|H)]P(H)}{P(A)} \\ &= \frac{(1 - 0.05) \times 0.002}{0.0518} \simeq 0.0367 = 3.76\%. \end{aligned}$$

Βλέπουμε, λοιπόν, πως η πιθανότητα ο εξεταζόμενος να είναι φορέας του HIV δεδομένου πως το αποτέλεσμα της εξέτασης είναι θετικό, είναι *σημαντικά μικρότερη* από την πρώτη μας «διαισθητική» απάντηση που ήταν βασισμένη στο σκεπτικό ότι κατά 95% η εξέταση δίνει το σωστό αποτέλεσμα! Η διαφορά αυτή προκύπτει από το ότι αρχικά δεν λάβαμε υπ' όψιν μας πως ο εξεταζόμενος επιλέχθηκε *τυχαία* από έναν πληθυσμό στον οποίο πολύ σπάνια συναντάμε έναν φορέα του HIV. Η αρχική πιθανότητα (2 στους χίλιους) να είναι φορέας φυσικά αυξάνεται (στο 3.76%) λόγω του θετικού αποτελέσματος της εξέτασης, αλλά δεν φτάνει ως το 95% όπως αρχικά φανταζόμασταν.

Ορισμός 3.2. (Διαμερίσεις) *Όταν ένα σύνολο ενδεχόμενων $B_1, B_2, \ldots, B_N$ είναι ξένα μεταξύ τους και επιπλέον $\bigcup_{i=1}^N B_i = \Omega$, τότε καλούνται* <u>διαμέριση</u> *του Ω.*

Λήμμα 3.4. (<u>Κανόνας Ολικής Πιθανότητας</u>) *Για οποιαδήποτε ενδεχόμενα A, B, όπου $P(B) > 0$ και $P(B') > 0$, ισχύει ότι*

$$P(A) = P(A|B)P(B) + P(A|B')P(B').$$

Γενικότερα, αν τα ενδεχόμενα $B_1, B_2, \ldots, B_N$ αποτελούν διαμέριση, με $P(B_i) > 0$ για όλα τα i, τότε:

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N).$$

Απόδειξη. Παρατηρούμε πως

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap (B_1 \cup B_2 \cup \cdots \cup B_N)) \\ &= P(AB_1 \cup AB_2 \cup \cdots \cup AB_N) \\ &= P(AB_1) + P(AB_2) + \cdots + P(AB_N) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N). \end{aligned}$$

Στην τρίτη ισότητα χρησιμοποιήσαμε την επιμεριστική ιδιότητα. Στην τέταρτη χρησιμοποιήσαμε το ότι έχουμε ένωση ξένων ενδεχόμενων. □

Παράδειγμα 3.10. Το 20% των Ελληνίδων είναι μελαχρινές, ενώ το 60% των Ελλήνων είναι μελαχρινοί. Αν επιλέξουμε ένα άτομο στην τύχη από το γενικό πληθυσμό, ποια είναι η πιθανότητα να είναι μελαχρινό; Διαισθητικά, η απάντηση είναι $\frac{1}{2} \times 20\% + \frac{1}{2} \times 60\% = 40\%$. Αυστηρά, μπορούμε να χρησιμοποιήσουμε τον νόμο της ολικής πιθανότητας. Έστω M το ενδεχόμενο ένα τυχαία επιλεγμένο άτομο να είναι μελαχρινό, και A το ενδεχόμενο να είναι άντρας. Μας έχει δοθεί ότι $P(M|A) = 0.6$, $P(M|A') = 0.2$. Αφού ο πληθυσμός είναι γενικός, υποθέτουμε πως $P(A) = 0.5$. Έχουμε:

$$P(M) = P(M|A)P(A) + P(M|A')P(A') = 0.6 \times 0.5 + 0.2 \times 0.5 = 0.4.$$

Αν η επιλογή του ατόμου ήταν από ένα στρατόπεδο με 90% ανδρικό πληθυσμό, και όχι από το γενικό πληθυσμό, τότε $P(A) = 0.9$ και ο άνω υπολογισμός γίνεται

$$P(M) = P(M|A)P(A) + P(M|A')P(A') = 0.6 \times 0.9 + 0.2 \times 0.1 = 0.56.$$

Λήμμα 3.5. (Κανόνας Bayes) *Για οποιαδήποτε ενδεχόμενα A, B με $P(A) > 0$, $P(B) > 0$, $P(B') > 0$ έχουμε:*

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B')P(B')}.$$

Γενικότερα, αν τα ενδεχόμενα $B_1, B_2, \ldots, B_N$ είναι διαμέριση, και $P(B_i) > 0$ για κάθε $i = 1, 2, \ldots, N$, και επιπλέον για το ενδεχόμενο A έχουμε $P(A) > 0$, τότε:

$$\begin{aligned} P(B_1|A) &= \frac{P(A|B_1)P(B_1)}{P(A)} \\ &= \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_N)P(B_N)}. \end{aligned}$$

Απόδειξη. Προφανής, χρησιμοποιώντας τον ορισμό της δεσμευμένης πιθανότητας και τον κανόνα της ολικής πιθανότητας. □

Παράδειγμα 3.11. Σε κάποιο πληθυσμό, 2% των ανθρώπων πάσχει από μια ασθένεια. Ένας τυχαία επιλεγμένος άνθρωπος κάνει ένα διαγνωστικό τεστ για αυτή την ασθένεια, όπου οι πιθανότητες σφάλματος του τεστ είναι:

$$P\big(\text{«αρνητικό τεστ»}\big|\text{«ασθενής»}\big) = 0.5\%,$$
$$P\big(\text{«θετικό τεστ»}\big|\text{«όχι ασθενής»}\big) = 5\%.$$

Δεδομένου ότι το τεστ είναι αρνητικό, ποια η πιθανότητα παρόλα αυτά να έχει την ασθένεια;

Ορίζουμε τα ενδεχόμενα

$$\begin{aligned} T &= \text{«θετικό τεστ»}, \\ A &= \text{«ασθενής»}, \end{aligned}$$

οπότε τα δεδομένα του προβλήματος είναι

$$P(A) = 0.02, \quad P(T'|A) = 0.005, \quad P(T|A') = 0.05,$$

και η ζητούμενη πιθανότητα είναι η $P(A|T')$. Από τον κανόνα του Bayes έχουμε

$$\begin{aligned} P(A|T') &= \frac{P(T'|A)P(A)}{P(T'|A)P(A) + P(T'|A')P(A')} \\ &= \frac{P(T'|A)P(A)}{P(T'|A)P(A) + [1 - P(T|A')][1 - P(A)]} \\ &= \frac{0.005 \times 0.02}{0.005 \times 0.02 + 0.95 \times 0.98} \simeq 10^{-4}. \end{aligned}$$

Παράδειγμα 3.12. Ανακοινώνεται από την εταιρία Dell ότι από τα PC ενός συγκεκριμένου μοντέλου που κυκλοφορούν στην αγορά:

1. Το 25% είναι ελαττωματικά,
2. το 25% δεν είναι ελαττωματικά, αλλά έχουν σημαντικό ρίσκο να παρουσιάσουν ελάττωμα,
3. και το 50% δεν έχουν κανένα από τα άνω δύο προβλήματα.

Διανέμεται από την εταιρία ένα πρόγραμμα διάγνωσης των προβλημάτων, το οποίο είναι 99% ακριβές, δηλαδή

$$P(T|(E \cup R)') = P(T'|E) = P(T'|R) = 1\%,$$

όπου τα ενδεχόμενα T, E, R, ορίζονται ως

$$\begin{aligned} T &= \text{«θετικό αποτέλεσμα διαγνωστικού τεστ»}, \\ E &= \text{«ελαττωματικό PC»}, \\ R &= \text{«PC με ρίσκο ελαττώματος»}. \end{aligned}$$

Κάνουμε στο PC μας αυτό το διαγνωστικό τεστ και βγαίνει αρνητικό. Ποια η πιθανότητα ο υπολογιστής μας πράγματι να μην έχει κανένα από τα δύο προβλήματα;

Εδώ ο κανόνας του Bayes μας δίνει

$$P((E \cup R)'|T') = \frac{P(T'|(E \cup R)')P((E \cup R)')}{P(T'|(E \cup R)')P((E \cup R)') + P(T'|E)P(E) + P(T'|R)P(R)},$$

όπου παρατηρούμε ότι τα $E, R, (E \cup R)'$ αποτελούν διαμέριση. Αντικαθιστώντας όπως και στο προηγούμενο παράδειγμα τις τιμές από τα δεδομένα του προβλήματος υπολογίζουμε

$$P((E \cup R)'|T') = \frac{0.99 \times 0.5}{0.99 \times 0.5 + 0.01 \times 0.25 + 0.01 \times 0.25} = 0.99.$$

Άρα, σε κάποιες (πολύ ειδικές) περιπτώσεις, μπορεί να έχουμε, για δύο ενδεχόμενα A, B, ότι $P(A|B) = P(B|A)$, όπως εδώ έχουμε $P((E \cup R)'|T') = P(T'|(E \cup R)')$.

Παράδειγμα 3.13. *(Τεστ πολλαπλών απαντήσεων)* Ένας μαθητής απαντά μια ερώτηση πολλαπλών απαντήσεων, με N δυνατές απαντήσεις. Αν ο μαθητής δεν ξέρει την απάντηση, επιλέγει μια από τις δυνατές απαντήσεις στην τύχη, χωρίς κάποια προτίμηση. Αν η πιθανότητα να ξέρει την απάντηση είναι p, ποια είναι η δεσμευμένη πιθανότητα να ήξερε την απάντηση δεδομένου ότι απάντησε σωστά;

Λύση: Ορίζουμε τα ενδεχόμενα A να απάντησε σωστά, και C ο μαθητής να γνώριζε την απάντηση, οπότε η ζητούμενη πιθανότητα είναι η $P(C|A)$. Έχουμε, από τον κανόνα του Bayes:

$$\begin{aligned} P(C|A) &= \frac{P(A|C)P(C)}{P(A|C)P(C) + P(A|C')P(C')} \\ &= \frac{1 \times p}{1 \times p + \frac{1}{N} \times (1-p)} = \frac{Np}{(N-1)p + 1}. \end{aligned}$$

Στα άνω, κάναμε την υπόθεση ότι αν ο μαθητής ξέρει την απάντηση, απαντά πάντα σωστά.

3.4 Ανεξαρτησία

Ορισμός 3.3. (Ανεξάρτητα ενδεχόμενα) *Δύο ενδεχόμενα A, B είναι ανεξάρτητα αν*

$$P(AB) = P(A)P(B).$$

Σε διαφορετική περίπτωση, καλούνται εξαρτημένα.

Παρατηρήσεις

1. Αν έχουμε $P(A) > 0$, τότε

$$P(AB) = P(A)P(B) \Leftrightarrow P(B) = \frac{P(AB)}{P(A)} = P(B|A).$$

2. Αντιστοίχως, αν $P(B) > 0$, τότε

$$P(AB) = P(A)P(B) \Leftrightarrow P(A) = \frac{P(AB)}{P(B)} = P(A|B).$$

3. Αν έχουμε $P(A) = 0$, τότε

$$AB \subseteq A \Rightarrow P(AB) \leq P(A) = 0 \Rightarrow P(AB) = 0,$$

   και επομένως τα A, B είναι ανεξάρτητα, αφού $P(AB) = P(A)P(B) = 0$.

4. Αν έχουμε $P(B) = 0$, ομοίως τα A, B είναι ανεξάρτητα.

Στις πρώτες δύο περιπτώσεις, που είναι οι πιο συνηθισμένες (αφού σπανίως μας ενδιαφέρουν ενδεχόμενα με πιθανότητα 0), δύο ενδεχόμενα είναι ανεξάρτητα αν, και μόνο αν, γνωρίζοντας αν συνέβη το ένα, δεν αλλάζει η πιθανότητα να συμβεί το άλλο.

Παράδειγμα 3.14. *(Ρίψεις ζαριού)* Ρίχνουμε ένα ζάρι δύο φορές. Υποθέτουμε ότι όλα τα αποτελέσματα είναι ισοπίθανα. Έστω τα ακόλουθα ενδεχόμενα:

1. $A =$ «Το άθροισμα είναι ζυγό»,
2. $B =$ «Το άθροισμα είναι μονό»,
3. $C =$ «Το άθροισμα είναι ίσο με 7»,
4. $D =$ «Φέραμε δύο φορές το ίδιο»,
5. $E =$ «Η πρώτη ζαριά ήρθε 6»,

6. $F =$ «Η πρώτη ζαριά ήρθε 1»,

7. $G =$ «Το άθροισμα είναι ίσο με 6».

Είναι ή όχι τα πιο κάτω ζεύγη ενδεχόμενων ανεξάρτητα;

1. A, B.

2. C, B.

3. D, E.

4. C, F.

5. G, F.

Ο χώρος πιθανότητας Ω αποτελείται από 36 αποτελέσματα, και συγκεκριμένα από τις διατεταγμένες δυάδες της μορφής (i, j), όπου $i, j = 1, \ldots, 6$. Εξ υποθέσεως, όλα τα αποτελέσματα έχουν την ίδια πιθανότητα, προφανώς $1/36$. Συνεπώς, μπορούμε να υπολογίσουμε την πιθανότητα οποιουδήποτε ενδεχόμενου, αρκεί να μετρήσουμε τον αριθμό των αποτελεσμάτων από τα οποία αποτελείται. Για παράδειγμα, $P(A) = P(B) = \frac{18}{36} = \frac{1}{2}$, αφού υπάρχουν 18 μονά και 18 ζυγά αποτελέσματα. Επιπλέον:

$$\begin{aligned} P(C) &= P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}) = \frac{6}{36} = \frac{1}{6}, \\ P(D) &= P(\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}) = \frac{6}{36} = \frac{1}{6}, \\ P(G) &= P(\{(1,5),(2,4),(3,3),(4,2),(5,1)\}) = \frac{5}{36}, \end{aligned}$$

ενώ προφανώς $P(E) = P(F) = 1/6$.

Για να αποφανθούμε για την ανεξαρτησία των ζητούμενων ζευγαριών, χρειαζόμαστε την πιθανότητα της τομής τους. Εύκολα, προκύπτει, αν μετρήσουμε τα αποτελέσματα που ανήκουν σε κάθε τομή, ότι:

$$\begin{aligned} P(A \cap B) &= P(\emptyset) = 0, \\ P(C \cap B) &= P(C) = \frac{1}{6}, \\ P(D \cap E) &= P((6,6)) = \frac{1}{36}, \\ P(C \cap F) &= P((1,6)) = \frac{1}{36}, \\ P(G \cap F) &= P((1,5)) = \frac{1}{36}. \end{aligned}$$

Έτσι, συγκρίνοντας την πιθανότητα της τομής με το γινόμενο των επί μέρους πιθανοτήτων σε κάθε περίπτωση, εύκολα προκύπτει ότι τα δύο ζεύγη D και E, C και F είναι ανεξάρτητα, ενώ τα άλλα όχι. Συμβαδίζουν τα αποτελέσματα αυτά με την διαίσθησή σας;

Παράδειγμα 3.15. *(Ανεξαρτησία συμπληρωμάτων)* Αν τα A και B είναι ανεξάρτητα ενδεχόμενα, τότε τα ζεύγη A και B', A' και B, και A' και B' είναι επίσης ανεξάρτητα. Πράγματι:

1. Το ότι τα A και B είναι ανεξάρτητα σημαίνει πως $P(A \cap B) = P(A)P(B)$. Από την ανεξαρτησία έχουμε

$$\begin{aligned} P(A)P(B') &= P(A)(1 - P(B)) \\ &= P(A) - P(A)P(B) = P(A) - P(A \cap B). \end{aligned}$$

   Επίσης

$$P(A) = P(A \cap B) + P(A \cap B')$$

   και συνδυάζοντας τις πιο πάνω σχέσεις έχουμε

$$P(A \cap B') = P(A)P(B'),$$

   συνεπώς τα A και B' είναι ανεξάρτητα.

2. Η ανεξαρτησία των A' και B προκύπτει όμοια με το προηγούμενο σκέλος.

3. Αφού τα A και B είναι ανεξάρτητα, θα είναι και τα A' και B, από το δεύτερο σκέλος. Εφαρμόζοντας τώρα το πρώτο σκέλος για τα A' και B, προκύπτει ότι θα είναι ανεξάρτητα και τα A' και B'.

Όλα τα αποτελέσματα έχουν διαισθητική εξήγηση. Για παράδειγμα, σχετικά με το πρώτο αποτέλεσμα, αν τα A, B είναι ανεξάρτητα, τότε αν μάθουμε ότι έγινε το B δεν αλλάζει η πιθανότητα να έχει συμβεί το A. Άρα, αν μάθουμε ότι έγινε το B', πάλι δεν θα αλλάξει η πιθανότητα του A.

Παράδειγμα 3.16. *(Δύο τμήματα)* Έστω δύο διαφορετικά τμήματα, εκ των οποίων το πρώτο αποτελείται από m_1 φοιτητές και w_1 φοιτήτριες, ενώ το δεύτερο αποτελείται από m_2 φοιτητές και w_2 φοιτήτριες. Επιλέγουμε ένα τμήμα στην τύχη, χωρίς κάποια προτίμηση, και εκ των υστέρων ένα άτομο από αυτό το τμήμα. Έστω A το ενδεχόμενο να επιλέξουμε το πρώτο τμήμα, και έστω M το ενδεχόμενο να επιλέξουμε φοιτητή. Ποια είναι η σχέση που πρέπει να ικανοποιούν τα $m_1,\ m_2,\ w_1,\ w_2$ ώστε να είναι ανεξάρτητα τα ενδεχόμενα A και M;

Εφόσον υποθέτουμε πως $P(A) = P(A') = \frac{1}{2}$, από τον κανόνα ολικής πιθανότητας,

$$P(M) = P(M|A)P(A) + P(M|A')P(A') = \frac{1}{2} \times \frac{m_1}{m_1 + w_1} + \frac{1}{2} \times \frac{m_2}{m_2 + w_2}.$$

Επιπλέον, από τον ορισμό της δεσμευμένης πιθανότητας,

$$P(M \cap A) = P(A)P(M|A) = \frac{1}{2} \times \frac{m_1}{m_1 + w_1}.$$

Άρα, για να έχουμε ανεξαρτησία των ενδεχόμενων M και A, εξ ορισμού απαιτείται:

$$\begin{aligned} & P(M \cap A) = P(M)P(A) \\ \Leftrightarrow\ & \frac{1}{2} \times \frac{m_1}{m_1 + w_1} = \left(\frac{1}{2} \times \frac{m_1}{m_1 + w_1} + \frac{1}{2} \times \frac{m_2}{m_2 + w_2}\right) \times \frac{1}{2} \\ \Leftrightarrow\ & \frac{m_1}{m_1 + w_1} = \frac{m_2}{m_2 + w_2} \\ \Leftrightarrow\ & \frac{m_1}{w_1} = \frac{m_2}{w_2}. \end{aligned}$$

Η απάντηση ήταν διαισθητικά αναμενόμενη: Οι αναλογίες φοιτητριών/φοιτητών σε κάθε τμήμα πρέπει να είναι ίδιες.

Παράδειγμα 3.17. Θα δείξουμε ότι αν το ενδεχόμενο A είναι ανεξάρτητο από τον εαυτό του, τότε αναγκαστικά θα έχουμε είτε $P(A) = 0$ ή $P(A) = 1$. Πράγματι, αν το A είναι ανεξάρτητο από τον εαυτό του, τότε η πιθανότητά του ικανοποιεί την

$$P(A) = P(A \cap A) = P(A)P(A) = (P(A))^2,$$

και οι μόνοι πραγματικοί αριθμοί x τέτοιοι ώστε $x^2 = x$ είναι το 0 και το 1.

Παράδειγμα 3.18. Αν $P(A) = 1$, τότε το A είναι ανεξάρτητο του B για οποιοδήποτε ενδεχόμενο B. Πράγματι, παρατηρήστε πως

$$P(A)P(B) = P(B) = P(BA \cup BA') = P(BA) + P(BA') = P(AB).$$

Η τρίτη ισότητα προκύπτει επειδή τα AB, $A'B$ είναι ξένα, ενώ η τέταρτη επειδή το $A'B$ είναι υποσύνολο του A' που έχει $P(A') = 0$, άρα και $P(A'B) = 0$.

Ορισμός 3.4. (Ανεξαρτησία πολλών ενδεχόμενων)

1. *Τρία ενδεχόμενα καλούνται <u>ανεξάρτητα</u> αν είναι ανά δύο ανεξάρτητα, δηλαδή*

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C),$$

και επιπλέον

$$P(ABC) = P(A)P(B)P(C).$$

2. n *ενδεχόμενα* $A_1, A_2, \ldots, A_n$ *καλούνται ανεξάρτητα αν η πιθανότητα της τομής οποιωνδήποτε από αυτά ισούται με το γινόμενο των αντίστοιχων πιθανοτήτων, δηλαδή*
$$P(A_{m(1)}A_{m(2)}\ldots A_{m(k)}) = P(A_{m(1)})P(A_{m(2)})\ldots P(A_{m(k)})$$
για κάθε υποσύνολο $\{A_{m(1)}, A_{m(2)}, \ldots, A_{m(k)}\}$ *των* n *ενδεχόμενων, όπου* $k \leq n$.

Παράδειγμα 3.19. Μπορεί τρία ενδεχόμενα να είναι ανά δύο ανεξάρτητα, χωρίς όμως να είναι (από κοινού) ανεξάρτητα. Για παράδειγμα, έστω $\Omega = \{1, 2, 3, 4\}$, και έστω $P(1) = P(2) = P(3) = P(4) = \frac{1}{4}$. Έστω επίσης τα ενδεχόμενα

$$A = \{1, 2\}, \qquad B = \{1, 3\}, \qquad C = \{1, 4\}.$$

Παρατηρήστε πως

$$\begin{aligned} P(A) = P(B) = P(C) &= \frac{1}{2}, \\ P(ABC) = P(AB) = P(BC) = P(AC) &= \frac{1}{4}, \\ P(AB) &= P(A)P(B), \\ P(AC) &= P(A)P(C), \\ P(BC) &= P(B)P(C), \\ P(ABC) = \frac{1}{4} &\neq \frac{1}{8} = P(A)P(B)P(C). \end{aligned}$$

Το αποτέλεσμα συμφωνεί με τη διαίσθησή μας. Αν γνωρίζουμε ότι συνέβη το B, δεν αλλάζουμε την πιθανότητα να συνέβη το A. Αν όμως γνωρίζουμε ότι συνέβη και το B και το C, είναι βέβαιο ότι συνέβη και το A.

Παρατηρήσεις

1. Σε περίπτωση που πρέπει να εξετάσουμε την ανεξαρτησία n ενδεχόμενων, πρέπει να ελέγξουμε $2^n - n - 1$ σχέσεις. Μπορείτε να εξηγήσετε γιατί;

2. Στα άνω παραδείγματα, η ανεξαρτησία (ή όχι) δύο (ή περισσοτέρων) ενδεχόμενων ήταν το *ζητούμενο.* Συνηθέστερα όμως, είναι η *αφετηρία* για υπολογισμό της πιθανότητας κάποιων ενδεχόμενων. Δηλαδή, χρησιμοποιούμε την ανεξαρτησία για να κατασκευάσουμε το μέτρο πιθανότητας, όπως ακριβώς χρησιμοποιήσαμε την υπόθεση ότι όλα τα αποτελέσματα είναι ισοπίθανα. Πρέπει βέβαια, η υπόθεση της ανεξαρτησίας να συμβαδίζει με την διαίσθησή μας. Δείτε τα επόμενα παραδείγματα.

Παράδειγμα 3.20. Έστω ότι στρίβουμε 4 φορές ένα κέρμα, το οποίο έρχεται κορώνα με πιθανότητα $P(H) = p$ και γράμματα με πιθανότητα $P(T) = 1-p$, για κάποιο (γνωστό

σε μας) $p \in (0,1)$. Εδώ ο χώρος πιθανότητας Ω αποτελείται από τα $2 \times 2 \times 2 \times 2 = 16$ δυνατά αποτελέσματα (π.χ., $HTHH$ ή $TTHH$, κλπ) των τεσσάρων ρίψεων. Αλλά αν το κέρμα δεν είναι δίκαιο (δηλαδή αν το $p \neq 1/2$) τα στοιχειώδη ενδεχόμενα *δεν* είναι ισοπίθανα οπότε δεν μπορούμε να χρησιμοποιήσουμε το Λήμμα 1.2.

Απ΄ την άλλη, λογικά μπορούμε να θεωρήσουμε ότι οι διαδοχικές ρίψεις είναι ανεξάρτητες μεταξύ τους. Δηλαδή, αν δύο ενδεχόμενα αφορούν δύο διαφορετικές ρίψεις, θα είναι ανεξάρτητα. Άρα μπορούμε να υπολογίσουμε την πιθανότητα ενός στοιχειώδους ενδεχομένου, όπως π.χ. του $\{HTHH\}$ ως εξής. Έστω H_1 το ενδεχόμενο η πρώτη ρίψη να είναι κορώνα, T_2 το ενδεχόμενο η δεύτερη ρίψη να είναι γράμματα, κ.ο.κ. Έχουμε:

$$\begin{aligned} P(HTHH) &= P\big(H_1 \cap T_2 \cap H_3 \cap H_4\big) = P(H_1) \times P(T_2) \times P(H_3) \times P(H_4) \\ &= p \times (1-p) \times p \times p = (1-p)p^3, \end{aligned}$$

όπου στη δεύτερη ισότητα θεωρούμε ότι τα τέσσερα πιο πάνω ενδεχόμενα είναι ανεξάρτητα.

Υπενθυμίζουμε ότι για να υπολογίσουμε την πιθανότητα *οποιουδήποτε* ενδεχομένου της μορφής $A = \{\omega_1, \omega_2, \ldots, \omega_k\}$, αρκεί να γνωρίζουμε την πιθανότητα των στοιχειωδών ενδεχομένων $\{\omega_i\}$, οπότε:

$$P(A) = P(\omega_1) + P(\omega_2) + \cdots + P(\omega_k).$$

Συνδυάζοντας αυτή την παρατήρηση με το γεγονός ότι στο παράδειγμά μας η ανεξαρτησία μας επιτρέπει να υπολογίσουμε την πιθανότητα για κάθε στοιχειώδες ενδεχόμενο, καταλήγουμε στο ότι μπορούμε να υπολογίσουμε την πιθανότητα *οποιουδήποτε ενδεχομένου* με χρήση της ανεξαρτησίας. Για παράδειγμα, αν $A =$«τέσσερις φορές το ίδιο αποτέλεσμα», τότε έχουμε:

$$\begin{aligned} P(A) &= P(\{HHHH, TTTT\}) = P\big(\{HHHH\} \cup \{TTTT\}\big) \\ &= P(HHHH) + P(TTTT) = p^4 + (1-p)^4. \end{aligned}$$

Παράδειγμα 3.21. Έστω ότι υπάρχουν τρεις συνδέσεις σε ένα δίκτυο, και η κάθε μία ενεργοποιείται, ανεξάρτητα από τις άλλες δύο, με πιθανότητα 1/4. Εξετάζουμε το ποιες συνδέσεις είναι ενεργές μια δεδομένη στιγμή, οπότε ο χώρος πιθανότητας Ω αποτελείται από τα 8 στοιχεία που περιγράφουν τις αντίστοιχες 8 δυνατές καταστάσεις. Για παράδειγμα, το στοιχείο EEA του Ω περιγράφει την κατάσταση όπου οι δύο πρώτες συνδέσεις είναι ενεργές και η τρίτη ανενεργή.

Ποια είναι η πιθανότητα του ενδεχομένου $B =$«ακριβώς μια σύνδεση είναι ενεργή»; Εφόσον τα στοιχειώδη ενδεχόμενα εδώ δεν είναι ισοπίθανα, θα ακολουθήσουμε την ίδια λογική όπως και στο προηγούμενο παράδειγμα:

$$\begin{aligned} P(B) &= P(\{EAA, AEA, AAE\}) = P\big(\{EAA\} \cup \{AEA\} \cup \{AAE\}\big) \\ &= P(EAA) + P(AEA) + P(AAE). \end{aligned}$$

Εφόσον η κάθε σύνδεση είναι ενεργή (με πιθανότητα 1/4) ή ανενεργή (με πιθανότητα 3/4) ανεξάρτητα από τις άλλες, έχουμε

$$P(B) = \frac{1}{4} \times \frac{3}{4} \times \frac{3}{4} + \frac{3}{4} \times \frac{1}{4} \times \frac{3}{4} + \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{27}{64}.$$

Παρατήρηση: Μερικές φορές καλούμαστε να αποδείξουμε, ή υποθέτουμε πως ισχύει, η ανεξαρτησία δύο ενδεχόμενων όταν είναι δεσμευμένα προς ένα τρίτο. Δηλαδή, έχουμε ως δεδομένο, ή ζητούμενο, την ακόλουθη σχέση:

$$P(A \cap B|C) = P(A|C)P(B|C).$$

Δείτε τα ακόλουθα παραδείγματα.

Παράδειγμα 3.22. *(Μεσογειακή αναιμία)* Σε κάποιο πληθυσμό, το 8% των ατόμων έχει το στίγμα της μεσογειακής αναιμίας. Έστω ότι ένα τυχαία επιλεγμένο άτομο κάνει μια εξέταση για να διαπιστώσει αν έχει το στίγμα ή όχι. Η εξέταση δεν είναι απόλυτα ακριβής, και η πιθανότητα το αποτέλεσμα να βγει θετικό ενώ δεν υπάρχει στίγμα είναι 10%. Επιπλέον, η πιθανότητα να βγει το αποτέλεσμα αρνητικό ενώ υπάρχει στίγμα είναι 1%. Ποιά η πιθανότητα να έχει στίγμα κάποιος που:

1. Κάνει την εξέταση και προκύπτει θετικό αποτέλεσμα;
2. Κάνει την εξέταση δύο ανεξάρτητες φορές, και προκύπτει θετικό αποτέλεσμα την πρώτη φορά και αρνητικό αποτέλεσμα την δεύτερη φορά;

Έχουμε, κατά περίπτωση:

1. Έστω A το ενδεχόμενο να βγει το αποτέλεσμα θετικό, και S να έχει το άτομο το στίγμα. Δίνεται ότι $P(S) = 0.08$, $P(A|S') = 0.1$ και $P(A'|S) = 0.01$. Έχουμε επίσης $P(S') = 0.92$ και $P(A|S) = 0.99$.

   Από το νόμο του Bayes, η ζητούμενη πιθανότητα $P(S|A)$ είναι:
   $$P(S|A) = \frac{P(A|S)P(S)}{P(A|S)P(S) + P(A|S')P(S')} = \frac{0.99 \times 0.08}{0.99 \times 0.08 + 0.1 \times 0.92} \simeq 0.4626.$$

2. Έστω A_1 το ενδεχόμενο να βγει το αποτέλεσμα θετικό στην πρώτη εξέταση, και A_2 το ενδεχόμενο να βγει το αποτέλεσμα θετικό στη δεύτερη εξέταση. Με εφαρμογή του νόμου του Bayes και της υπόθεσης της ανεξαρτησίας, έχουμε:
   $$\begin{aligned} P(S|A_1 \cap A_2') &= \frac{P(A_1 \cap A_2'|S)P(S)}{P(A_1 \cap A_2'|S)P(S) + P(A_1 \cap A_2'|S')P(S')} \\ &= \frac{P(A_1|S)P(A_2'|S)P(S)}{P(A_1|S)P(A_2'|S)P(S) + P(A_1|S')P(A_2'|S')P(S')} \\ &= \frac{0.99 \times 0.01 \times 0.08}{0.99 \times 0.01 \times 0.08 + 0.1 \times 0.9 \times 0.92} \simeq 0.0095. \end{aligned}$$

Παρατηρήστε ότι στη δεύτερη ισότητα χρησιμοποιήσαμε το ότι τα αποτελέσματα της εξέτασης είναι ανεξάρτητα, δεδομένου ότι το επιλεγμένο άτομο έχει (ή δεν έχει) στίγμα.

Παράδειγμα 3.23. *(Δύο ζάρια)* Έστω ότι έχουμε τρία ζάρια, όπου το πρώτο είναι δίκαιο, το δεύτερο δεν έρχεται ποτέ 6, και το τρίτο έχει $P(6) = 2/3$. Επιλέγουμε ένα στην τύχη, χωρίς κάποια προτίμηση, το ρίχνουμε δύο φορές και φέρνουμε 66. Ποια η πιθανότητα να επιλέξαμε το δίκαιο ζάρι;

Ορίζουμε τα ενδεχόμενα F, A_2, A_3 ως «επιλέξαμε το δίκαιο ζάρι», «επιλέξαμε το δεύτερο ζάρι» και «επιλέξαμε το τρίτο ζάρι», αντίστοιχα. Επιπλέον ορίζουμε τα ενδεχόμενα E_1 =«το πρώτο ζάρι ήταν 6», E_2 =«το δεύτερο ζάρι ήταν 6», $E = E_1E_2$=«φέραμε 66». Από τον κανόνα του Bayes,

$$\begin{aligned} P(F|E) &= \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|A_2)P(A_2) + P(E|A_3)P(A_3)} \\ &= \frac{P(E_1E_2|F)P(F)}{P(E_1E_2|F)P(F) + P(E_1E_2|A_2)P(A_2) + P(E_1E_2|A_3)P(A_3)} \\ &= \frac{(1/6)(1/6)(1/3)}{(1/6)(1/6)(1/3) + 0 \times (1/3) + (2/3)(2/3)(1/3)} = \frac{1}{16}, \end{aligned}$$

όπου στην τρίτη ισότητα υποθέσαμε ανεξαρτησία μεταξύ των ρίψεων, *με δεδομένη την επιλογή κάποιου ζαριού*, συνεπώς, για παράδειγμα, $P(E_1E_2|F) = P(E_1|F)P(E_2|F)$.

Πράγματι, εφόσον φέραμε 2 φορές 6, είναι πιθανότερο να επιλέξαμε το ζάρι το οποίο έχει πιθανότητα 2/3 να φέρει 6, παρά το δίκαιο ζάρι.

Σαν τελευταία παρατήρηση, εύκολα μπορούμε να δούμε πως, χωρίς δέσμευση, τα E_1, E_2 δεν είναι ανεξάρτητα, δηλαδή $P(E_1E_2) \neq P(E_1)P(E_2)$. Πράγματι, αν φέρουμε 6 στην πρώτη ρίψη, τότε αποκλείεται να επιλέξαμε το δεύτερο ζάρι, και συνεπώς αυξάνει η πιθανότητα του ενδεχόμενου να φέρουμε 6 στη δεύτερη ρίψη.

Παράδειγμα 3.24. *(Εταιρία πληροφορικής — συνέχεια)* Επιστρέφουμε στην εταιρεία πληροφορικής του Παραδείγματος 3.1, στην οποία η κατανομή των εργαζόμενων είναι η ακόλουθη:

	Έλληνες	αλλοδαποί
τεχνικοί	22	25
προγραμματιστές	18	5

Όπως και στο Παράδειγμα 3.1, επιλέγουμε κάποιον από τους 70 εργαζόμενους, χωρίς κάποια προτίμηση.

Δίνεται, επιπλέον, ότι η κατανομή ανά φύλλο έχει ως εξής:

(Άντρες)	Έλληνες	αλλοδαποί
τεχνικοί	17	20
προγραμματιστές	13	0

(Γυναίκες)	Έλληνες	αλλοδαποί
τεχνικοί	5	5
προγραμματιστές	5	5

Ορίζουμε τα ενδεχόμενα

$$T = «\text{επελέγη τεχνικός}», \quad E = «\text{επελέγη Έλληνας}», \quad K = «\text{επελέγη Γυναίκα}».$$

Παρατηρήστε πως τα T, E δεν είναι ανεξάρτητα, διότι

$$P(T) = \frac{47}{70}, \quad P(E) = \frac{4}{7}, \quad P(ET) = \frac{22}{70},$$

και δεν ισχύει ότι $P(ET) = P(E)P(T)$. Όμως, έχουμε

$$P(T|K) = \frac{1}{2}, \quad P(E|K) = \frac{1}{2}, \quad P(ET|K) = \frac{1}{4},$$

επομένως $P(ET|K) = P(E|K)P(T|K)$ και τα ενδεχόμενα E, T είναι ανεξάρτητα αν τα δεσμεύσουμε ως προς το K. Βεβαιωθείτε ότι καταλαβαίνετε διαισθητικά γιατί ισχύει το αποτέλεσμα.

Ορισμός 3.5. (Ανεξαρτησία υποπειραμάτων) *Δύο ή περισσότερα υποπειράματα ενός σύνθετου πειράματος καλούνται ανεξάρτητα αν οποιοδήποτε σύνολο από ενδεχόμενα, κάθε ένα εκ των οποίων αφορά ένα σύνολο υποπειραμάτων (από τα άνω) ξένα από τα αντίστοιχα σύνολα υποπειραμάτων των άλλων ενδεχόμενων, είναι ανεξάρτητα μεταξύ τους.*

Παρατηρήσεις

1. Ο άνω ορισμός δεν είναι απολύτως αυστηρός, αλλά για τα πλαίσια αυτού του μαθήματος επαρκεί. Ο ορισμός φαίνεται πολύπλοκος, αλλά η εφαρμογή του είναι απλή.

2. Μερικά παραδείγματα:

   (α΄) Αν δύο ζαριές είναι ανεξάρτητες, τα ενδεχόμενα να έρθει το πρώτο ζάρι 1 και το δεύτερο ζάρι να μην έρθει 4 είναι ανεξάρτητα.

   (β΄) Αν κάνουμε 10 παιδιά διαδοχικά και οι γεννήσεις είναι ανεξάρτητες, τότε το ενδεχόμενο το 5ο παιδί να είναι αγόρι είναι ανεξάρτητο από το ενδεχόμενο να έχουμε κάνει αγόρι στις πρώτες 4 γεννήσεις.

   (γ΄) Αν ρίξουμε 5 πέναλτι διαδοχικά, και οι εκτελέσεις είναι ανεξάρτητες, το ενδεχόμενο να βάλουμε το πρώτο είναι ανεξάρτητο από το ενδεχόμενο να βάλουμε κάποιο από τα 4 υπόλοιπα.

   (δ΄) Μπορείτε να βρείτε σε ποια από τα παραδείγματα αυτού του κεφαλαίου εμφανίζονται ανεξάρτητα υποπειράματα;

Κεφάλαιο 4

Διακριτές Τυχαίες Μεταβλητές

Σε σύνθετα προβλήματα πιθανοτήτων, όπως π.χ. σε προβλήματα ανάλυσης πολύπλοκων δικτύων ή αλγορίθμων, η λεπτομερής, στοιχείο-προς-στοιχείο περιγραφή του πλήρους δειγματικού χώρου Ω είναι εξαιρετικά χρονοβόρα, και συχνά είναι και περιττή. Όπως θα δούμε σε αυτό το κεφάλαιο, η έννοια της *τυχαίας μεταβλητής* μας επιτρέπει να απλοποιήσουμε ή και να παρακάμψουμε αυτή τη διαδικασία.

4.1 Διακριτές Τυχαίες Μεταβλητές

Ορισμός 4.1. (Διακριτές τυχαίες μεταβλητές)

1. *Διακριτή Τυχαία Μεταβλητή (Τ.Μ.) καλείται κάθε συνάρτηση $X : \Omega \to \mathbb{R}$ όπου ο Ω είναι ένας δειγματικός χώρος, και για την οποία το σύνολο S των τιμών που μπορεί να λάβει η X είναι πεπερασμένο ή αριθμήσιμο.*

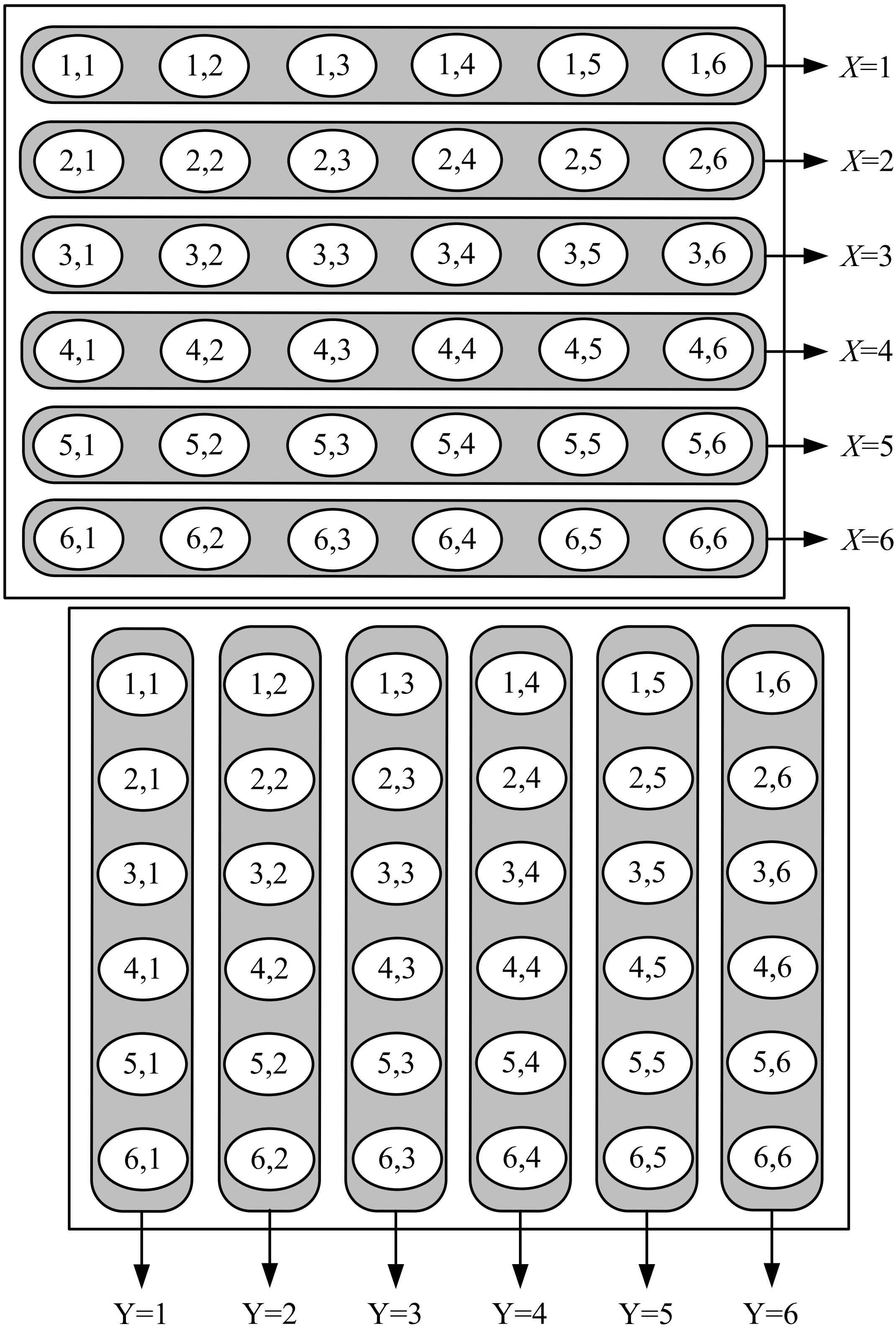
2. *Το σύνολο S καλείται σύνολο τιμών της X.*

3. *Η συνάρτηση μάζας πιθανότητας, ή μάζα πιθανότητας, ή απλώς μάζα μιας διακριτής X είναι η συνάρτηση $p : S \to [0, 1]$ η οποία ορίζεται ως*

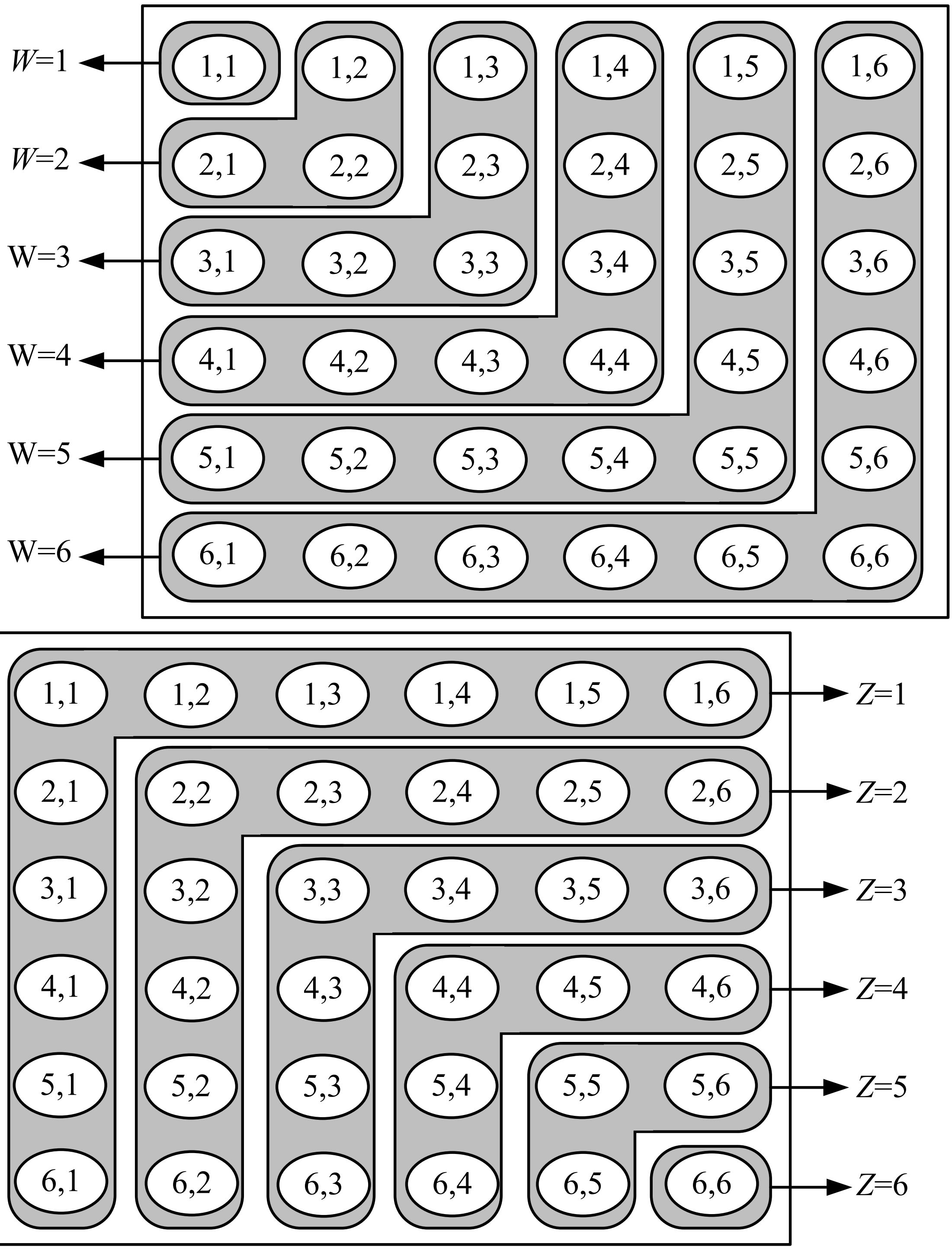
$$p(x) = P(X = x) = P(\{\omega : X(\omega) = x\}), \quad \forall x \in S.$$

4. *Η συνάρτηση κατανομής πιθανότητας, ή κατανομή πιθανότητας, ή απλώς κατανομή της X είναι η συνάρτηση $F : \mathbb{R} \to [0, 1]$ η οποία ορίζεται ως*

$$F(x) = P(X \le x) = P(\{\omega : X(\omega) \le x\}), \quad \forall x \in \mathbb{R}.$$



Σχήμα 4.1: Οι Τ.Μ. X, Y του Παραδείγματος 4.1.



Σχήμα 4.2: Οι Τ.Μ. W, Z του Παραδείγματος 4.1.

Παρατηρήσεις

1. Εκτός από τις διακριτές, υπάρχουν και άλλα είδη τυχαίων μεταβλητών. Προς το παρόν, όταν μιλάμε για μια Τ.Μ., θα εννοείται πως αυτή είναι διακριτή.

2. Αν υποθέσουμε κάποιο μέτρο πιθανότητας για τον δειγματικό χώρο, εύκολα μπορούμε να βρούμε τις πιθανότητες με τις οποίες λαμβάνουν τις διάφορες τιμές τους οι Τ.Μ που ορίζονται σε αυτόν, και να υπολογίσουμε έτσι την μάζα και την κατανομή τους. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 4.1. *(Δύο ζάρια)* Ρίχνουμε διαδοχικά δύο ζάρια, και, όπως συνήθως, ορίζουμε τον δειγματικό χώρο ως εξής:

$$\Omega = \{11, 12, \ldots, 16, 21, 22, \ldots, 26, \ldots, 61, 62, \ldots, 66\}.$$

Έστω X και Y τα αποτελέσματα των δύο ρίψεων. Επιπλέον, ορίζουμε τις Τ.Μ. $W = \max\{X, Y\}$ και $Z = \min\{X, Y\}$. Στα Σχήματα 4.1, 4.2, έχουμε δείξει, για κάθε μία απ′ τις Τ.Μ. X, Y, W, Z, και για όλες τις τιμές που μπορεί να πάρουν, το σύνολο όλων των αποτελεσμάτων που προκαλούν αυτή την τιμή. Προφανώς, τα σύνολα τιμών $S_X = S_Y = S_W = S_Z = \{1, 2, 3, 4, 5, 6\}$.

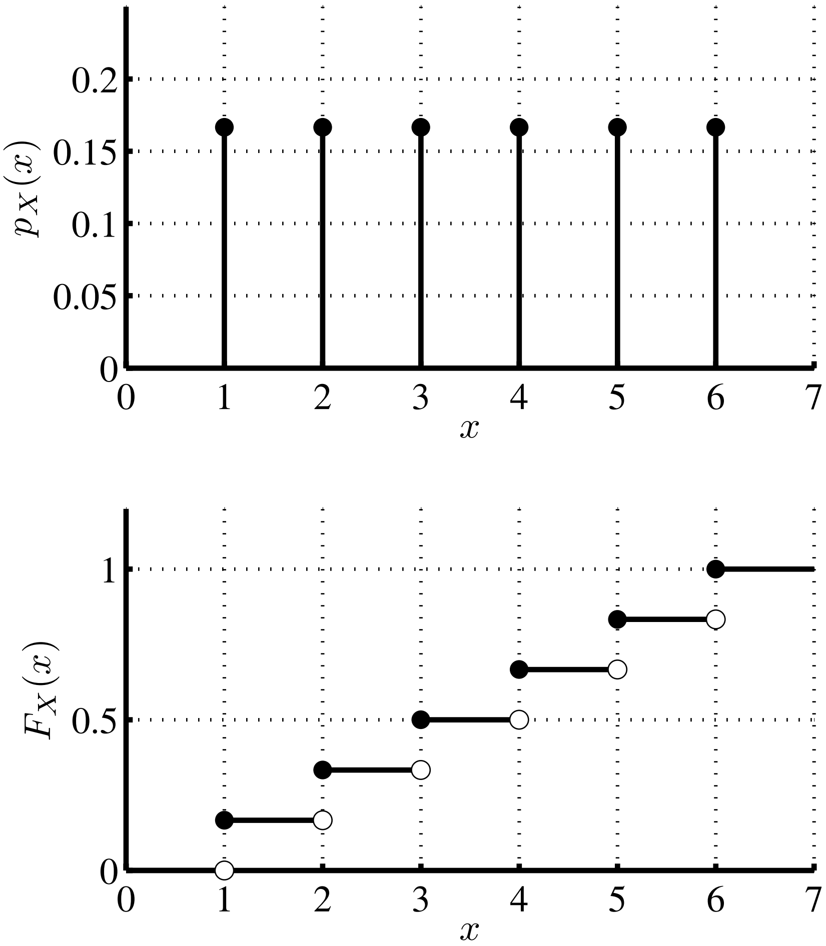
Αν τώρα υποθέσουμε ότι όλα τα αποτελέσματα είναι ισοπίθανα, τότε έχουμε ότι

$$\begin{aligned}
P(X = 1) &= P(\{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6)\}) \\
&= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{6}, \\
P(Y = 3) &= P(\{(1,3), (2,3), (3,3), (4,3), (5,3), (6,3)\}) \\
&= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{6}, \\
P(W = 3) &= P(\{(1,3), (2,3), (3,3), (3,2), (3,1)\}) \\
&= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{5}{36}, \\
P(Z = 5) &= P(\{(6,5), (5,5), (5,6)\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{12}.
\end{aligned}$$

Παρόμοια θα μπορούσαμε να υπολογίσουμε τις άνω πιθανότητες ακόμα και αν το μέτρο πιθανότητας του δειγματικού χώρου ήταν διαφορετικό, για παράδειγμα αν κάποιοι συνδυασμοί ρίψεων ήταν πιο πιθανοί από άλλους. Θα παραμείνουμε όμως στην περίπτωση που όλα τα αποτελέσματα είναι ισοπίθανα, και θα υπολογίσουμε τις μάζες και τις κατανομές των X, Y, W, Z.

Για κάθε δυνατή τιμή $x = 1, 2, 3, 4, 5, 6$ της X υπάρχουν 6 στοιχεία του Ω που αντιστοιχούν σε αυτή την τιμή, και άρα η πιθανότητα του ενδεχομένου $\{X = x\}$ είναι $6/36 = 1/6$. Συνεπώς η μάζα της X είναι η

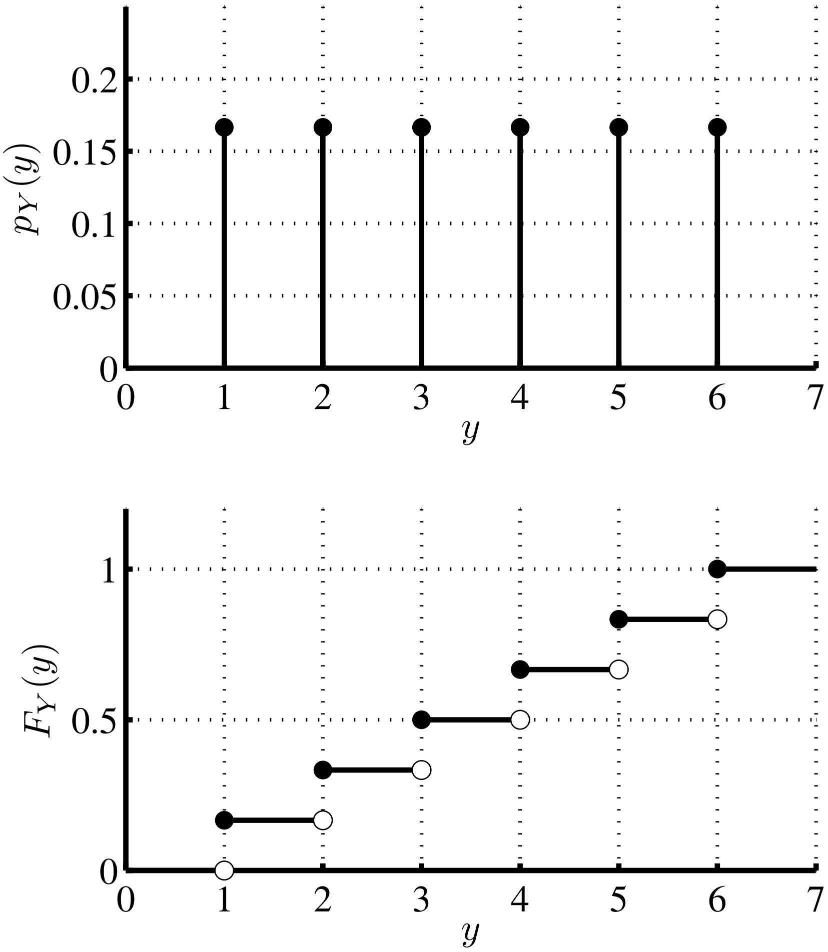
$$p_X(x) = \frac{1}{6}, \quad x = 1, 2, \ldots, 6.$$



Σχήμα 4.3: Η μάζα και η συνάρτηση κατανομής της Τ.Μ. X στο Παράδειγμα 4.1.

Η συνάρτηση κατανομής $F_X(x)$ της X εύκολα μπορεί να υπολογιστεί, με παρόμοιο τρόπο, αλλά λαμβάνοντας υπ' όψιν ότι το όρισμα της είναι πραγματικός αριθμός:

1. Αν $x < 1$, τότε προφανώς $F_X(x) = P(X \leq x) = 0$.
2. Αν $x \in [1, 2)$, τότε $F_X(x) = P(X \leq x) = P(X = 1) = \frac{1}{6}$.
3. Αν $x = [2, 3)$, τότε $F_X(x) = P(X \leq x) = P(X = 1) + P(X = 2) = \frac{2}{6}$.
4. Αν $x = [3, 4)$, τότε $F_X(x) = P(X \leq x) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{3}{6}$.
5. Αν $x = [4, 5)$, τότε $F_X(x) = P(X \leq x) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = \frac{4}{6}$.
6. Αν $x = [5, 6)$, τότε $F_X(x) = P(X \leq x) = P(X = 1) + \cdots + P(X = 5) = \frac{5}{6}$.
7. Αν $x \geq 6$, τότε $F_X(x) = P(X \leq x) = P(X = 1) + \cdots + P(X = 6) = 1$.

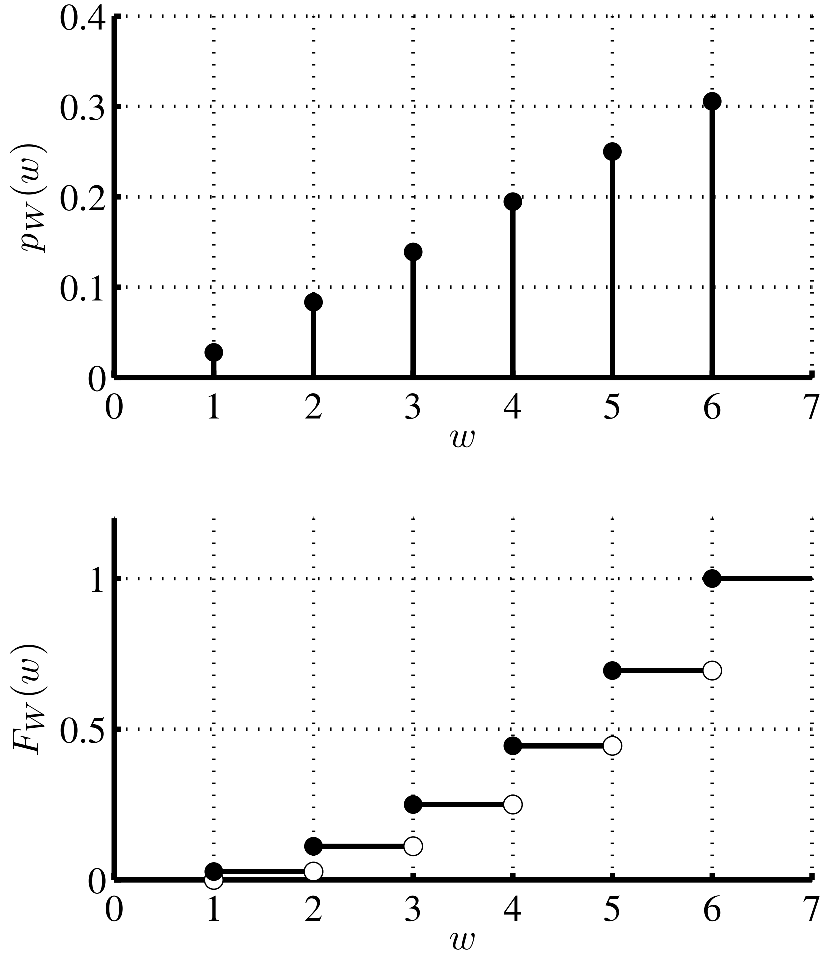


Σχήμα 4.4: Η μάζα και η συνάρτηση κατανομής της Τ.Μ. Y στο Παράδειγμα 4.1.

Προκύπτει τελικά πως

$$F_X(x) = \begin{cases} 0, & x < 1, \\ 1/6, & 1 \leq x < 2, \\ 2/6, & 2 \leq x < 3, \\ 3/6, & 3 \leq x < 4, \\ 4/6, & 4 \leq x < 5, \\ 5/6, & 5 \leq x < 6, \\ 1, & 6 \leq x. \end{cases}$$

Παρατηρήστε ότι στις τιμές $1, 2, 3, 4, 5, 6$ η κατανομή παρουσιάζει άλμα, και επιπλέον ότι είναι παντού δεξιά συνεχής.



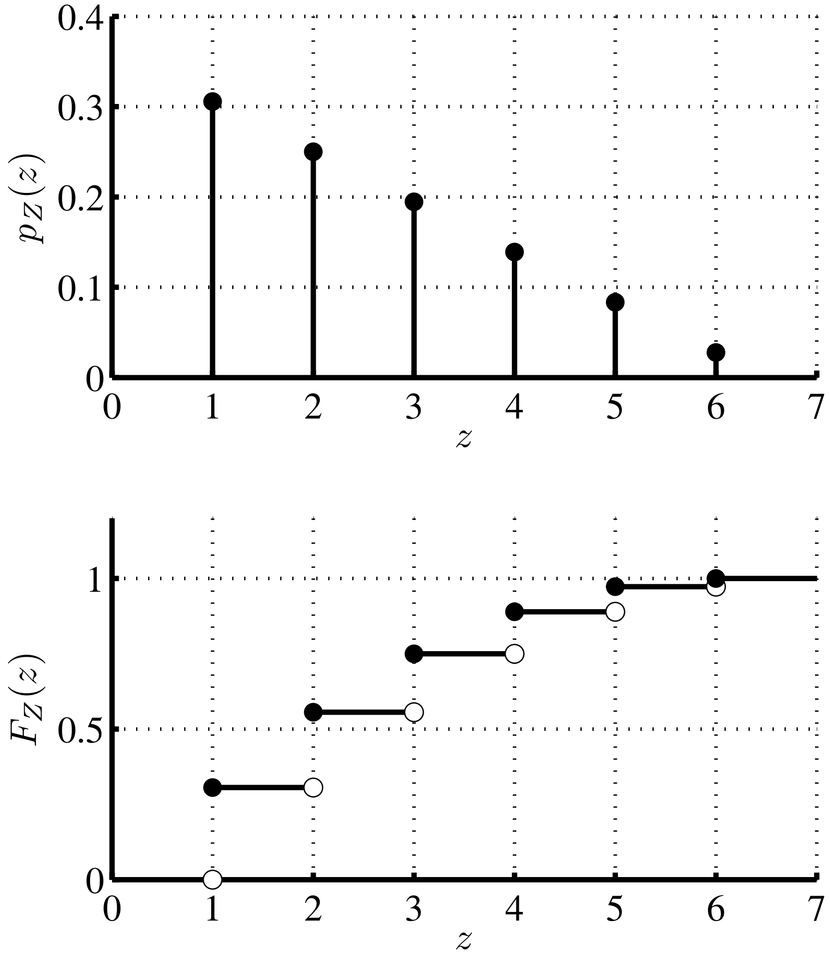
Σχήμα 4.5: Η μάζα και η συνάρτηση κατανομής της Τ.Μ. W στο Παράδειγμα 4.1.

Με ακριβώς τον ίδιο τρόπο βρίσκουμε πως η Y έχει την ίδια μάζα και την ίδια συνάρτηση κατανομής με την X. Δηλαδή

$$p_Y(y) = \frac{1}{6}, \quad y = 1, 2, \ldots, 6,$$

και

$$F_Y(y) = \begin{cases} 0, & y < 1, \\ 1/6, & 1 \leq y < 2, \\ 2/6, & 2 \leq y < 3, \\ 3/6, & 3 \leq y < 4, \\ 4/6, & 4 \leq y < 5, \\ 5/6, & 5 \leq y < 6, \\ 1, & 6 \leq y. \end{cases}$$



Σχήμα 4.6: Η μάζα και η συνάρτηση κατανομής της Τ.Μ. Z στο Παράδειγμα 4.1.

Η μάζα της $W = \max\{X_1, X_2\}$ υπολογίζεται όπως και της X, μετρώντας τα αποτελέσματα του αρχικού Δ.Χ. που αντιστοιχούν στις διάφορες τιμές w που μπορεί να πάρει. Με τη βοήθεια του Σχήματος 4.2 βρίσκουμε:

$$p_W(1) = \frac{1}{36},\ p_W(2) = \frac{3}{36},\ p_W(3) = \frac{5}{36},\ p_W(4) = \frac{7}{36},\ p_W(5) = \frac{9}{36},\ p_W(6) = \frac{11}{36},$$

$$F_W(w) = \begin{cases} 0, & w < 1, \\ 1/36, & 1 \leq w < 2, \\ 4/36, & 2 \leq w < 3, \\ 9/36, & 3 \leq w < 4, \\ 16/36, & 4 \leq w < 5, \\ 25/36, & 5 \leq w < 6, \\ 1, & 6 \leq w. \end{cases}$$

Τέλος, η μάζα της $Z = \min\{X_1, X_2\}$ υπολογίζεται εύκολα με την ίδια μεθοδολογία.

Με τη βοήθεια του Σχήματος 4.2 βρίσκουμε:

$$p_Z(1) = \frac{11}{36},\ p_Z(2) = \frac{9}{36},\ p_Z(3) = \frac{7}{36},\ p_Z(4) = \frac{5}{36},\ p_Z(5) = \frac{3}{36},\ p_Z(6) = \frac{1}{36}.$$

$$F_Z(z) = \begin{cases} 0, & z < 1, \\ 11/36, & 1 \leq z < 2, \\ 20/36, & 2 \leq z < 3, \\ 27/36, & 3 \leq z < 4, \\ 32/36, & 4 \leq z < 5, \\ 35/36, & 5 \leq z < 6, \\ 1, & 6 \leq z. \end{cases}$$

Οι μάζες και οι κατανομές των X, Y, W, Z έχουν σχεδιαστεί στα Σχήματα 4.3, 4.4, 4.5, και 4.6.

Παρατηρήσεις

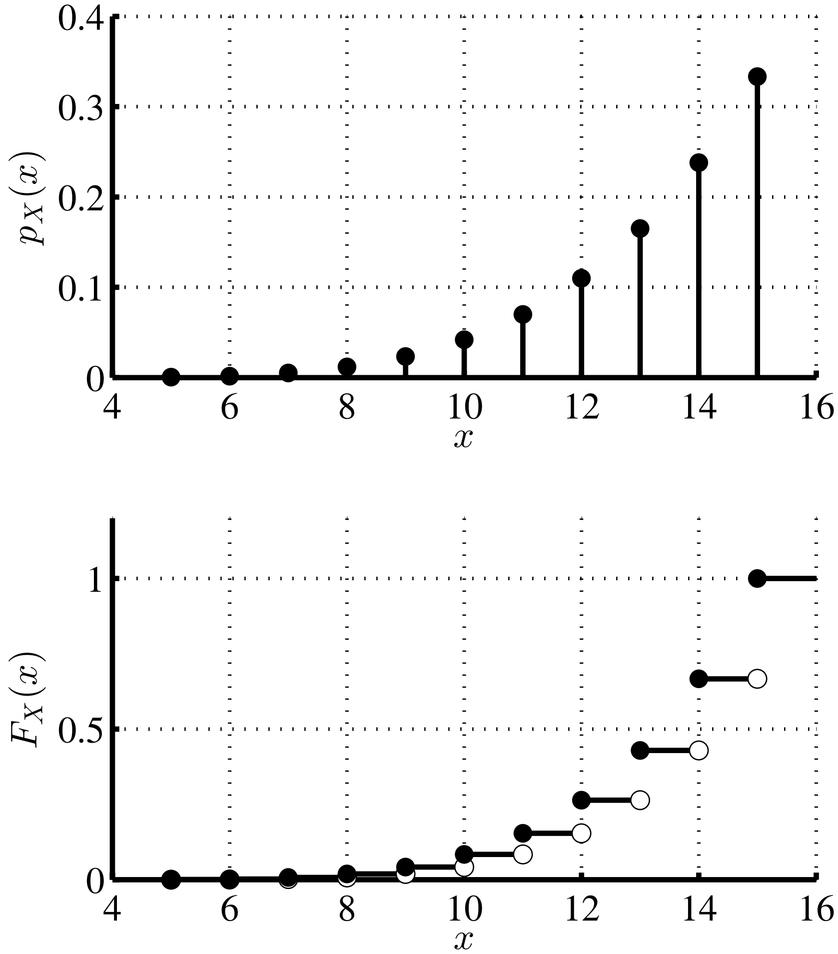
1. Αν και δεν το δείχνουμε εδώ, μπορούμε να αποδείξουμε αυστηρά ότι ο συνδυασμός του μέτρου πιθανότητας στον δειγματικό χώρο Ω και η συνάρτηση X δημιουργούν *ένα νέο μέτρο πιθανότητας*, αυτή τη φορά στο σύνολο τιμών S του X. Η $P(A)$ στο νέο μέτρο πιθανότητας ισούται με την $P(\{\omega : X(\omega) \in A\})$ στο αρχικό μέτρο πιθανότητας. Δεν θα χρησιμοποιούμε διαφορετικά σύμβολα (π.χ., $P_1(\cdot)$ και $P_2(\cdot)$) για τα δύο διαφορετικά μέτρα πιθανότητας.

2. Αν ένα μέτρο $P(\cdot)$ επάγεται από μια Τ.Μ., θα γράφουμε την $P(A)$ και ως $P(X \in A)$. Για παράδειγμα, $P((-2, 3]) = P(-2 < X \leq 3)$.

3. Μερικές φορές η μάζα καλείται και πυκνότητα, προκειμένου να υπάρχει ομοιομορφία με την περίπτωση των συνεχών Τ.Μ., που θα εξετάσουμε σε επόμενα κεφάλαια.

4. Παρατηρήστε πως, όταν χρειάζεται, χρησιμοποιούμε τον αντίστοιχο δείκτη για να ξεχωρίζουμε τα σύνολα τιμών, τις μάζες και τις κατανομές διαφόρων Τ.Μ. Για παράδειγμα, η X έχει σύνολο τιμών S_X, μάζα $p_X(x)$ και κατανομή $F_X(x)$, κ.ο.κ.

Παράδειγμα 4.2. *(Επιτροπή βουλής)* Μια επιτροπή της βουλής, αποτελούμενη από 15 άτομα, δημιουργεί μια υποεπιτροπή, με τον ακόλουθο τρόπο: επιλέγεται ένας συνδυασμός 5 ατόμων, ανάμεσα στα 15, χωρίς κάποια προτίμηση στο συνδυασμό, και από τα επιλεγμένα άτομα ο αρχαιότερος ορίζεται πρόεδρος της υποεπιτροπής, και ο δεύτερος αρχαιότερος ορίζεται γραμματέας της υποεπιτροπής. Έστω $1, 2, \ldots, 15$ τα άτομα της επιτροπής, σύμφωνα με την αρχαιότητά τους (ο νεότερος είναι ο 1 και ο αρχαιότερος είναι ο 15).



Σχήμα 4.7: Η μάζα και η συνάρτηση κατανομής της Τ.Μ. X της Άσκησης 4.2.

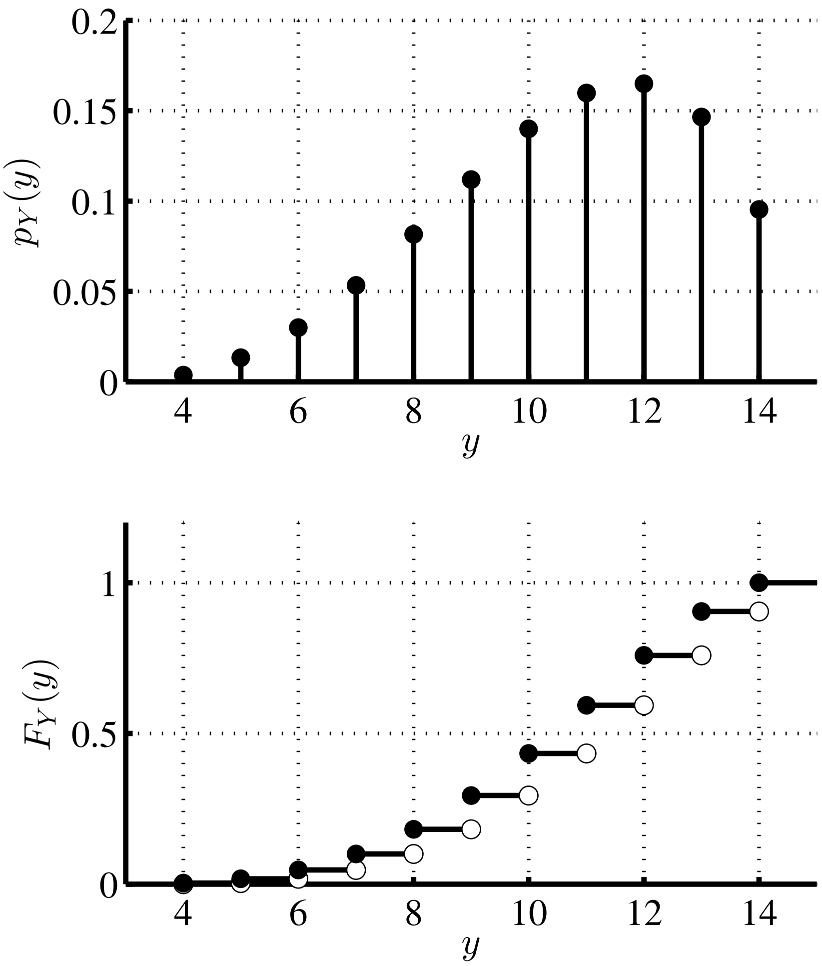
Έστω X και Y Τ.Μ. που δηλώνουν το άτομο που θα είναι πρόεδρος και γραμματέας της υποεπιτροπής αντίστοιχα. Θα υπολογίσουμε τις μάζες και τις κατανομές των X, Y.

Καταρχήν, παρατηρούμε πως ο δειγματικός χώρος Ω του πειράματος έχει μέγεθος $|\Omega| = \binom{15}{5}$. Επίσης, $S_X = \{5, 6, \ldots, 15\}$ και $S_Y = \{4, 5, \ldots 14\}$.

Σχετικά με την μάζα του X, παρατηρούμε πως $X = x$, όπου $x = 5, 6, \ldots, 15$, αν και μόνο αν επιλεγούν για την υποεπιτροπή το άτομο x και άλλα 4 άτομα αποκλειστικά στα πρώτα $x - 1$. Αυτό μπορεί να γίνει με $\binom{x-1}{4}$ τρόπους. Άρα,

$$p_X(x) = P(X = x) = \binom{x-1}{4} \Big/ \binom{15}{5}, \quad x = 5, 6, \ldots, 15.$$

Σχετικά με την κατανομή του X, προφανώς $F_X(x) = 0$ για $x < 5$ και $F_X(x) = 1$ για $x > 15$. Στην περίπτωση $5 \leq x \leq 15$, παρατηρούμε πως $X \leq x$ αν και μόνο αν επιλεγούν για την υποεπιτροπή άτομα αποκλειστικά ανάμεσα στα πρώτα $\lfloor x \rfloor$. Αυτό



Σχήμα 4.8: Η μάζα και η συνάρτηση κατανομής της Τ.Μ. Y της Άσκησης 4.2.

μπορεί να γίνει με $\binom{\lfloor x \rfloor}{5}$ τρόπους. Άρα,

$$F_X(x) = P(X \le x) = \binom{\lfloor x \rfloor}{5} \Big/ \binom{15}{5}, \quad 5 \le x \le 15.$$

Σχετικά με την μάζα του Y, παρατηρούμε πως $Y = y$, όπου $y = 4, 5, \ldots, 14$, αν και μόνο αν επιλεγούν για την υποεπιτροπή κάποιο από τα άτομα $y+1, y+2, \ldots, 15$, το άτομο y και άλλα 3 άτομα αποκλειστικά στα πρώτα $y-1$. Αυτό μπορεί να γίνει με $(15-y) \times \binom{y-1}{3}$ τρόπους. Άρα,

$$p_Y(y) = P(Y = y) = (15-y)\binom{y-1}{3} \Big/ \binom{15}{5}, \quad y = 4, 5, \ldots, 14.$$

Σχετικά με την κατανομή του Y, καταρχήν $F_Y(y) = 0$ αν $y < 4$ και $F_Y(y) = 1$ αν $y > 14$. Για την περίπτωση $4 \le y \le 14$, παρατηρούμε πως $Y \le y$ αν και μόνο αν γίνει ένα από τα ακόλουθα δύο ενδεχόμενα:

1. επιλεγούν για την υποεπιτροπή άτομα αποκλειστικά ανάμεσα στα πρώτα $\lfloor y \rfloor$. Αυτό μπορεί να γίνει με $\binom{\lfloor y \rfloor}{5}$ τρόπους. (Αν $y \leq 4 < 5$, τότε $\binom{\lfloor y \rfloor}{5} = 0$.)

2. επιλεγούν για την υποεπιτροπή 4 άτομα αποκλειστικά ανάμεσα στα πρώτα $\lfloor y \rfloor$, και ένα άτομο ανάμεσα στα $\lfloor y \rfloor + 1, \lfloor y \rfloor + 2, \ldots, 15$. Αυτό μπορεί να γίνει με $\binom{\lfloor y \rfloor}{4} \times (15 - \lfloor y \rfloor)$ τρόπους.

Άρα,

$$F_Y(y) = P(Y \leq y) = \frac{\binom{\lfloor y \rfloor}{5} + \binom{\lfloor y \rfloor}{4} \times (15 - \lfloor y \rfloor)}{\binom{15}{5}}, \quad 4 \leq y \leq 14.$$

Η μάζα και η κατανομή του X έχουν σχεδιαστεί στο Σχήμα 4.7 ενώ η μάζα και η κατανομή του Y έχουν σχεδιαστεί στο Σχήμα 4.8.

Παρατηρήσεις

1. Στο άνω παράδειγμα, λόγω του πλήθους των αποτελεσμάτων, δεν σχεδιάσαμε τον δειγματικό χώρο. Είναι πάντως πολύ σημαντικό να θυμόμαστε ότι για να υπολογίσουμε την μάζα και την κατανομή μιας Τ.Μ., καθώς και την πιθανότητα οποιουδήποτε ενδεχόμενου εμπλέκει την Τ.Μ., είναι απαραίτητο να πάμε στον δειγματικό χώρο Ω και να προσδιορίσουμε εκεί την πιθανότητα κάποιου κατάλληλου ενδεχόμενου $A \in \Omega$.

2. Παρακάτω παρατίθενται οι βασικές ιδιότητες της μάζας και της κατανομής. Επαληθεύστε τις για τα άνω παραδείγματα.

Λήμμα 4.1. (Ιδιότητες κατανομής) *Η κατανομή $F(x)$ έχει τις εξής ιδιότητες:*

1. *Η $F(x)$ είναι αύξουσα.*

2. *Η $F(x)$ είναι δεξιά συνεχής.*

3. $P(X < a) = F(a^-) \triangleq \lim_{x \to a^-} F(x)$.

4. $\lim_{x \to \infty} F(x) = 1$ *και* $\lim_{x \to -\infty} F(x) = 0$.

Απόδειξη. Η $F(x)$ είναι αύξουσα γιατί

$$x_1 \leq x_2 \Rightarrow \{\omega : \ X(\omega) \leq x_1\} \subseteq \{\omega : \ X(\omega) \leq x_2\} \Rightarrow P(X \leq x_1) \leq P(X \leq x_2).$$

Οι αποδείξεις των υπόλοιπων ιδιοτήτων παραλείπονται λόγω της δυσκολίας τους. □

Λήμμα 4.2. (Ιδιότητες μάζας) *Έστω διακριτή Τ.Μ. X με σύνολο τιμών $S = \{x_1, x_2, \ldots\}$. Η μάζα $p(x)$ έχει τις εξής ιδιότητες:*

1. $0 \le p(x) \le 1$ *για κάθε* $x \in S$.
2. $\sum_{x \in S} p(x) = 1$.
3. $P(X \in T) = \sum_{x \in T \cap S} p(x)$ *για κάθε* $T \subseteq \mathbb{R}$.

Απόδειξη. Η πρώτη ιδιότητα είναι προφανής (η $p(x)$ είναι πιθανότητα). Επίσης,

$$\sum_{x \in S} p(x) = \sum_{i=1}^{\infty} P(X = x_i) = P(\cup_{i=1}^{\infty}\{X = x_i\}) = 1.$$

Η δεύτερη ισότητα προέκυψε γιατί τα $\{X = x_i\}$ είναι ξένα, ενώ η τρίτη γιατί όλα μαζί απαρτίζουν τον δειγματικό χώρο Ω. Η τελευταία ιδιότητα προκύπτει ως εξής:

$$\begin{aligned} P(X \in T) &= P(\cup_{x \in T \cap S}\{X = x\}) + P(\cup_{x \in T \cap S'}\{X = x\}) \\ &= \sum_{x \in T \cap S} P(X = x) + 0 = \sum_{x \in T \cap S} p(x). \end{aligned}$$

Στη δεύτερη ισότητα χρησιμοποιούμε το ότι τα ενδεχόμενα της ένωσης είναι ξένα. □

Λήμμα 4.3. (Σχέση μεταξύ μάζας και κατανομής) *Έστω διακριτή Τ.Μ. X με σύνολο τιμών $S = \{x_1, x_2, \ldots\}$.*

1. *Αν διαθέτουμε την κατανομή $F(x)$, η μάζα $p(x)$ προκύπτει ως εξής:*

$$p(x) = F(x) - F(x^-), \quad \forall x \in S.$$

2. *Αν διαθέτουμε την μάζα $p(x)$, η κατανομή $F(x)$ προκύπτει ως εξής:*

$$F(x) = \sum_{y \in S:\ y \le x} p(y), \quad \forall x \in \mathbb{R}.$$

Απόδειξη. 1.

$$\begin{aligned} P(X \le x) &= P(\{X < x\} \cup \{X = x\}) = P(X < x) + P(X = x) \\ &\Rightarrow p(x) = P(X = x) = P(X \le x) - P(X < x) = F(x) - F(x^-). \end{aligned}$$

Στην τελευταία ισότητα χρησιμοποιήσαμε το Λήμμα 4.1.

2. Παρατηρούμε πως $F(x) = P(X \le x)$ άρα μπορούμε να χρησιμοποιήσουμε την τρίτη ιδιότητα του Λήμματος 4.2, και προκύπτει άμεσα το ζητούμενο.

□

4.2 Μέση Τιμή

Παράδειγμα 4.3. *(N ζαριές)* Έστω πως ρίχνουμε N ανεξάρτητες φορές ένα δίκαιο ζάρι. Περίπου πόσος περιμένουμε ότι θα είναι ο μέσος όρος των ζαριών; Σκεφτόμαστε ως εξής: Από τις N, περίπου μια στις 6, δηλαδή $N/6$, θα είναι 1, περίπου άλλες $N/6$ θα είναι 2, κ.ο.κ., και τελικά ο μέσος όρος θα είναι περίπου

$$\frac{1}{N}\left(1\times\frac{N}{6}+\cdots+6\times\frac{N}{6}\right)=1\times\frac{1}{6}+\cdots+6\times\frac{1}{6}=\frac{7}{2}.$$

Ο συγκεκριμένος υπολογισμός εμφανίζεται γενικευμένος στον επόμενο ορισμό.

Ορισμός 4.2. (Μέση τιμή) *Έστω διακριτή Τ.Μ. X με σύνολο τιμών S και μάζα $p(x)$. Ορίζουμε ως τη μέση τιμή ή αναμενόμενη τιμή ή προσδοκώμενη τιμή της X το άθροισμα*

$$E(X)\triangleq\sum_{x\in S}xp(x).$$

Παρατηρήσεις

1. Συχνά χρησιμοποιείται τα σύμβολο μ για την μέση τιμή μιας Τ.Μ.

2. Διαισθητικά, όπως φαίνεται και από το άνω παράδειγμα, περιμένουμε ότι ο μέσος όρος πολλών Τ.Μ. κατανεμημένων όπως η Τ.Μ. X θα είναι κοντά στη μέση τιμή $E(X)$. Η διαίσθησή μας θα επαληθευτεί όταν μιλήσουμε αργότερα για τον Νόμο των Μεγάλων Αριθμών και το Κεντρικό Οριακό Θεώρημα.

3. Διαισθητικά, επίσης, περιμένουμε ότι οι τιμές διαδοχικών Τ.Μ. κατανεμημένων όπως η X «τείνουν να κυμαίνονται» γύρω από την $E(X)$.

4. Στην συνηθισμένη περίπτωση που η Τ.Μ. X παίρνει τιμές στους μη αρνητικούς ακέραιους, έχουμε
$$E(X)=\sum_{x=0}^{\infty}xp(x).$$

5. Αν, ακόμα πιο απλά, η Τ.Μ. X παίρνει τιμές στο σύνολο $\{0,1,\ldots,N\}$, έχουμε
$$E(X)=\sum_{x=0}^{N}xp(x).$$

6. Στην ειδική περίπτωση που η Τ.Μ. X λαμβάνει μόνο μια τιμή, δηλαδή $S=\{c\}$, $p(c)=1$, ο ορισμός συμφωνεί με τη διαίσθησή μας, καθώς δίνει $E(X)=c\times 1=c$.

Παράδειγμα 4.4. *(Δύο ζάρια — συνέχεια)* Για τις Τ.Μ. X, Y, W, Z του Παραδείγματος 4.1, έχουμε:

$$\begin{aligned} E(X) = E(Y) &= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2}, \\ E(W) &= 1 \times \frac{1}{36} + 2 \times \frac{3}{36} + 3 \times \frac{5}{36} + 4 \times \frac{7}{36} + 5 \times \frac{9}{36} + 6 \times \frac{11}{36} = \frac{161}{36}, \\ E(Z) &= 1 \times \frac{11}{36} + 2 \times \frac{9}{36} + 3 \times \frac{7}{36} + 4 \times \frac{5}{36} + 5 \times \frac{3}{36} + 6 \times \frac{1}{36} = \frac{91}{36}. \end{aligned}$$

Λήμμα 4.4. (Μέση τιμή συνάρτησης Τ.Μ.) *Έστω X διακριτή Τ.Μ. με σύνολο τιμών $S_X = \{x_1, x_2, \dots\}$ και μάζα $p_X(x)$.*

1. *Έστω $Y = g(X)$ μια νέα διακριτή Τ.Μ., συνάρτηση της προηγούμενης. Θα ισχύει:*
$$E(Y) = \sum_{x \in S_X} g(x) p_X(x).$$
2. *Αν $Y = aX + b$, όπου $a, b \in \mathbb{R}$, έχουμε $E(aX + b) = aE(X) + b$.*
3. *Πιο γενικά, αν $Y = a_1 g_1(X) + a_2 g_2(X) + \cdots + a_n g_n(X) = \sum_{i=1}^{n} a_i g_i(X)$, τότε*
$$E\left(\sum_{i=1}^{n} a_i g_i(X)\right) = \sum_{i=1}^{n} a_i E(g_i(X)).$$

Απόδειξη. 1. Παρατηρούμε πως:

$$\begin{aligned} E(Y) &= \sum_{y \in S_Y} y p_Y(y) = \sum_{y \in S_Y} y \left(\sum_{x : g(x) = y} p_X(x) \right) \\ &= \sum_{y \in S_Y} \sum_{x : g(x) = y} g(x) p_X(x) = \sum_{x \in S_X} g(x) p_X(x). \end{aligned}$$

2. Παρατηρούμε πως:

$$\begin{aligned} E(aX + b) &= \sum_{x \in S_X} (ax + b) p(x) = \sum_{x \in S_X} axp(x) + \sum_{x \in S_X} bp(x) \\ &= a \sum_{x \in S_X} xp(x) + b \sum_{x \in S_X} p(x) = aE(X) + b. \end{aligned}$$

3. Πιο γενικά, έχουμε:

$$\begin{aligned} E\left(\sum_{i=1}^{n} a_i g_i(X)\right) &= \sum_{x\in S_X}\left(\sum_{i=1}^{n} a_i g_i(x)\right)p(x) = \sum_{i=1}^{n}\sum_{x\in S_X} a_i g_i(x)p(x) \\ &= \sum_{i=1}^{n} a_i \sum_{x\in S_X} g_i(x)p(x) = \sum_{i=1}^{n} a_i E(g_i(X)). \end{aligned}$$

□

Παρατηρήσεις

1. Στην συνηθισμένη περίπτωση που η Τ.Μ. X παίρνει τιμές στους μη αρνητικούς ακέραιους, έχουμε
$$E(g(X)) = \sum_{x=0}^{\infty} g(x)p(x).$$

2. Αν, ακόμα πιο απλά, η Τ.Μ. X παίρνει τιμές στο σύνολο $\{0, 1, \ldots, N\}$, έχουμε
$$E(g(X)) = \sum_{x=0}^{N} g(x)p(x).$$

Παράδειγμα 4.5. Ένας στρατιώτης εκτελεί μια βολή με το όπλο του, και το οριζόντιο σφάλμα που κάνει σε μέτρα, X, είναι μια Τ.Μ. με μάζα

$$p_X(x) = \begin{cases} 0.1, & x = -2, \\ 0.2, & x = -1, \\ 0.4, & x = 0, \\ 0.2, & x = 1, \\ 0.1, & x = 2. \end{cases}$$

Δηλαδή, με πιθανότητα 0.1 η βολή θα καταλήξει 2 μέτρα αριστερά, με πιθανότητα 0.2 θα καταλήξει 1 μέτρο δεξιά, με πιθανότητα 0.4 θα καταλήξει στο στόχο, κ.ο.κ. Ανάλογα με το αποτέλεσμα, ο λοχαγός τιμωρεί το στρατιώτη με $Y = X^2$ ημέρες φυλάκισης. Κατά μέσο όρο, πόσες μέρες τιμωρία παίρνει ο στρατιώτης ανά βολή; Δηλαδή, πόσο είναι το $E(Y) = E(X^2)$;

Για να απαντήσουμε το ερώτημα, έχουμε δύο επιλογές. Η πρώτη είναι να υπολογίσουμε την μάζα της Τ.Μ. Y και ακολούθως την μέση τιμή της, από τον ορισμό της

μέσης τιμής. Παρατηρήστε πως

$$\begin{aligned} p_Y(0) &= P(Y=0) = P(X=0) = 0.4, \\ p_Y(1) &= P(Y=1) = P(X=1) + P(X=-1) = 0.2 + 0.2 = 0.4, \\ p_Y(4) &= P(Y=4) = P(X=2) + P(X=-2) = 0.1 + 0.1 = 0.2. \end{aligned}$$

Άρα,

$$E(Y) = 0 \times 0.4 + 1 \times 0.4 + 4 \times 0.2 = 1.2.$$

Η δεύτερη επιλογή είναι να χρησιμοποιήσουμε το πρώτο σκέλος του Λήμματος 4.4, που παρακάμπτει τον υπολογισμό της μάζας της Y:

$$E(Y) = \sum_{x \in S_X} x^2 p_X(x) = (-2)^2 \times 0.1 + (-1)^2 \times 0.2 + 0^2 \times 0.4 + 1^2 \times 0.2 + 2^2 \times 0.1 = 1.2.$$

Όπως μπορείτε να παρατηρήσετε, η δεύτερη μέθοδος είναι πιο άμεση. Υπάρχουν μάλιστα πολλές περιπτώσεις όπου ο υπολογισμός της μάζας της Y είναι τόσο δύσκολος, που πρακτικά μόνο η δεύτερη μέθοδος μπορεί να χρησιμοποιηθεί.

Παράδειγμα 4.6. *(Δύο ζάρια — συνέχεια)* Έστω το ακόλουθο παιχνίδι. Ρίχνουμε δύο ζάρια, και λαμβάνουμε ως κέρδος 10 φορές το μέγιστο $W = \max\{X, Y\}$ των δύο ζαριών X, Y σε Ευρώ. Για να παίξουμε όμως το παιχνίδι, πρέπει να καταβάλουμε στην αρχή 40 ευρώ. Μας συμφέρει να παίξουμε αυτό το παιχνίδι; Για να απαντήσουμε, παρατηρούμε πως το καθαρό μας κέρδος είναι $C = 10W - 40$, άρα πρέπει να υπολογίσουμε την $E(C) = E(10W - 40)$.

Όπως και στο προηγούμενο παράδειγμα, μπορούμε να ακολουθήσουμε δύο τρόπους. Ο πρώτος είναι να υπολογίσουμε την μάζα της C, και από εκεί την $E(C)$ μέσω του ορισμού της μέσης τιμής. Έχουμε:

$$\begin{aligned} p_C(-30) &= P(C=-30) = P(W=1) = \frac{1}{36}, \\ p_C(-20) &= P(C=-20) = P(W=2) = \frac{3}{36}, \\ p_C(-10) &= P(C=-10) = P(W=3) = \frac{5}{36}, \\ p_C(0) &= P(C=0) = P(W=4) = \frac{7}{36}, \\ p_C(10) &= P(C=10) = P(W=5) = \frac{9}{36}, \\ p_C(20) &= P(C=20) = P(W=6) = \frac{11}{36}. \end{aligned}$$

Άρα,

$$E(C) = (-30) \times \frac{1}{36} + (-20) \times \frac{3}{36} + (-10) \times \frac{5}{36} + 0 \times \frac{7}{36} + 10 \times \frac{9}{36} + 20 \times \frac{11}{36} = \frac{85}{18},$$

και τελικά μας συμφέρει να παίξουμε.

Ο δεύτερος τρόπος είναι να χρησιμοποιήσουμε το δεύτερο σκέλος του Λήμματος 4.4, που παρακάμπτει τον υπολογισμό της μάζας της C:

$$E(C) = E(10W - 40) = 10E(W) - 40 = 10 \times \frac{161}{36} - 40 = \frac{85}{18}.$$

Όπως μπορείτε να δείτε, ο δεύτερος τρόπος είναι πολύ πιο άμεσος.

Παράδειγμα 4.7. *(Άπειρη μέση τιμή)* Έστω μια Τ.Μ. X με σύνολο τιμών το $S = \mathbb{N} = \{1, 2, \dots\}$ και μάζα

$$p(k) = \frac{C}{k^2}, \quad k \geq 1.$$

Το άθροισμα όλων των τιμών της πρέπει να ισούται με 1, συνεπώς πρέπει να έχουμε

$$1 = \sum_{k=1}^{\infty} p(k) = \sum_{k=1}^{\infty} \frac{C}{k^2} = C \sum_{k=1}^{\infty} \frac{1}{k^2} = C \times \frac{\pi^2}{6},$$

όπου χρησιμοποιήσαμε την προφανή ισότητα $\sum_{k=1}^{\infty}(1/k^2) = \pi^2/6$. Άρα, $C = 6/\pi^2 = 0.6079271\dots$.

Η μέση τιμή της X είναι

$$E(X) = \sum_{k=1}^{\infty} k\, p(k) = \sum_{k=1}^{\infty} k\, \frac{6}{\pi^2} \frac{1}{k^2} = \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k},$$

το οποίο ισούται με ∞ αφού, ως γνωστόν, η αρμονική σειρά $\sum_{k=1}^{\infty}(1/k)$ αποκλίνει. Άρα, βρήκαμε μια Τ.Μ. με άπειρη μέση τιμή!

4.3 Διασπορά

Ορισμός 4.3. (Διασπορά) *Έστω διακριτή Τ.Μ. X με μέση τιμή $E(X)$. Η διασπορά της ορίζεται ως*

$$\mathrm{VAR}(X) = E\big[(X - E(X))^2\big].$$

Η τυπική απόκλιση της X ορίζεται ως $\sqrt{\mathrm{VAR}(X)}$.

Παρατηρήσεις

1. Χρησιμοποιούμε επίσης το σύμβολο σ^2 για τη διασπορά, και συνεπώς το σ για την τυπική απόκλιση.

2. Η διασπορά εκφράζει πόσο μεγάλο είναι το «άπλωμα», γύρω από τη μέση τιμή, διαδοχικών Τ.Μ. με κατανομή σαν της X. Σαν παράδειγμα, σκεφτείτε πόση είναι η διασπορά μιας τετριμμένης Τ.Μ. που λαμβάνει μια μόνο τιμή με πιθανότητα 1. Παρατηρήστε επίσης ότι η διασπορά της Τ.Μ. X με μάζα $p_X(0) = 1/2$, $p_X(a) = 1/2$ είναι

$$\mathrm{VAR}(X) = E\big[(X - E(X))^2\big] = E\left[\left(X - \frac{a}{2}\right)^2\right] = \frac{1}{2}\left(\frac{a}{2}\right)^2 + \frac{1}{2}\left(\frac{a}{2}\right)^2 = \frac{a^2}{4}.$$

Παράδειγμα 4.8. *(Δύο ζάρια — συνέχεια)* Για τις Τ.Μ. X, Y, W, Z του Παραδείγματος 4.1, έχουμε:

$$\begin{aligned}
\mathrm{VAR}(X) &= E\left[(X - E(X))^2\right] \\
&= \left(1 - \frac{7}{2}\right)^2 \times \frac{1}{6} + \left(2 - \frac{7}{2}\right)^2 \times \frac{1}{6} + \left(3 - \frac{7}{2}\right)^2 \times \frac{1}{6} \\
&\quad + \left(4 - \frac{7}{2}\right)^2 \times \frac{1}{6} + \left(5 - \frac{7}{2}\right)^2 \times \frac{1}{6} + \left(6 - \frac{7}{2}\right)^2 \times \frac{1}{6} = \frac{35}{12}, \\
\mathrm{VAR}(Y) &= \mathrm{VAR}(X) = \frac{35}{12}, \\
\mathrm{VAR}(W) &= E\left[(W - E(W))^2\right] \\
&= \left(1 - \frac{161}{36}\right)^2 \times \frac{1}{36} + \left(2 - \frac{161}{36}\right)^2 \times \frac{3}{36} + \left(3 - \frac{161}{36}\right)^2 \times \frac{5}{36} \\
&\quad + \left(4 - \frac{161}{36}\right)^2 \times \frac{7}{36} + \left(5 - \frac{161}{36}\right)^2 \times \frac{9}{36} + \left(6 - \frac{161}{36}\right)^2 \times \frac{11}{36} \\
&= \frac{2555}{1296},
\end{aligned}$$

$$\begin{aligned}
\mathrm{VAR}(Z) &= E\left[(Z-E(Z))^2\right] \\
&= \left(1-\frac{91}{36}\right)^2 \times \frac{11}{36} + \left(2-\frac{91}{36}\right)^2 \times \frac{9}{36} + \left(3-\frac{91}{36}\right)^2 \times \frac{7}{36} \\
&\quad + \left(4-\frac{91}{36}\right)^2 \times \frac{5}{36} + \left(5-\frac{91}{36}\right)^2 \times \frac{3}{36} + \left(6-\frac{91}{36}\right)^2 \times \frac{1}{36} = \frac{2555}{1296}.
\end{aligned}$$

Μια ματιά στα Σχήματα 4.5, 4.6 ίσως σας πείσει πως πράγματι, είναι λογικό τα W, Z να έχουν ακριβώς την ίδια διασπορά.

Λήμμα 4.5. (Ιδιότητες διασποράς) *Έστω διακριτή Τ.Μ. X. Ισχύουν τα εξής:*

$$\begin{aligned}
\mathrm{VAR}(X) &= E(X^2) - (E(X))^2, \\
\mathrm{VAR}(aX+b) &= a^2\mathrm{VAR}(X), \quad \forall a, b \in \mathbb{R}.
\end{aligned}$$

Απόδειξη.

$$\begin{aligned}
\mathrm{VAR}(X) &= E\big[(X-E(X))^2\big] = E[X^2 + (E(X))^2 - 2XE(X)] \\
&= E(X^2) + (E(X))^2 - 2E(E(X)X) \\
&= E(X^2) + (E(X))^2 - 2(E(X))^2 = E(X^2) - (E(X))^2, \\
\mathrm{VAR}(aX+b) &= E((aX+b)^2) - (E(aX+b))^2 \\
&= E(a^2X^2 + b^2 + 2abX) - (aE(X)+b)^2 \\
&= a^2E(X^2) + b^2 + 2abE(X) - a^2(E(X))^2 - b^2 - 2abE(X) \\
&= a^2(E(X^2) - (E(X))^2) = a^2\mathrm{VAR}(X).
\end{aligned}$$

□

Παράδειγμα 4.9. *(Δύο ζάρια — συνέχεια)* Για τις Τ.Μ. X, Y, W, Z του Παραδείγματος 4.1, έχουμε:

$$\begin{aligned}
E(X^2) = E(Y^2) &= 1^2 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{6} \\
&\quad + 4^2 \times \frac{1}{6} + 5^2 \times \frac{1}{6} + 6^2 \times \frac{1}{6} = \frac{91}{6}, \\
E(W^2) &= 1^2 \times \frac{1}{36} + 2^2 \times \frac{3}{36} + 3^2 \times \frac{5}{36} \\
&\quad + 4^2 \times \frac{7}{36} + 5^2 \times \frac{9}{36} + 6^2 \times \frac{11}{36} = \frac{791}{36}, \\
E(Z^2) &= 1^2 \times \frac{11}{36} + 2^2 \times \frac{9}{36} + 3^2 \times \frac{7}{36} \\
&\quad + 4^2 \times \frac{5}{36} + 5^2 \times \frac{3}{36} + 6^2 \times \frac{1}{36} = \frac{301}{36}.
\end{aligned}$$

Συνεπώς,

$$\begin{aligned}\mathrm{VAR}(X) = \mathrm{VAR}(Y) &= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12},\\ \mathrm{VAR}(W) &= \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \frac{2555}{1296},\\ \mathrm{VAR}(Z) &= \frac{301}{36} - \left(\frac{91}{36}\right)^2 = \frac{2555}{1296}.\end{aligned}$$

Παρατηρήστε ότι με χρήση του Λήμματος 4.5 μπορέσαμε να υπολογίσουμε τις διασπορές των X, Y, W, Z με κάπως λιγότερες πράξεις από ότι με χρήση του ορισμού της διασποράς.

Παράδειγμα 4.10. *(Η μέθοδος της δεύτερης ροπής)* Έστω μια Τ.Μ. X με σύνολο τιμών $S = \{0, 1, 2, \dots\}$, με μέση τιμή μ και διασπορά σ^2. Αν η τυπική απόκλιση σ της X είναι σημαντικά μικρότερη από τη μέση τιμή μ, τότε η πιθανότητα το X να ισούται με μηδέν είναι μικρή. Συγκεκριμένα, ισχύει ότι

$$P(X = 0) \leq \left(\frac{\sigma}{\mu}\right)^2.$$

Πράγματι, έστω $p(k)$, για $k \in S = \{0, 1, 2, \dots\}$ η μάζα της X. Από τον ορισμό της διασποράς έχουμε

$$\sigma^2 = \sum_{k=0}^{\infty} (k - \mu)^2 p(k),$$

και εφόσον όλοι οι όροι του πιο πάνω αθροίσματος είναι μεγαλύτεροι ή ίσοι του μηδενός, ολόκληρο το άθροισμα θα είναι μεγαλύτερο ή ίσο του όρου $k = 0$, και τελικά

$$\sigma^2 = \sum_{k=0}^{\infty} (k - \mu)^2 p(k) \geq (0 - \mu)^2 p(0) = \mu^2 P(X = 0),$$

που είναι ακριβώς η ζητούμενη ανισότητα.

4.4 Παραδείγματα

Παράδειγμα 4.11. *(Τυχερό παιχνίδι)* Σε ένα παιχνίδι, ο παίκτης έχει τυχαίο κέρδος $X = -20, 0, 10$ Ευρώ, με αντίστοιχες πιθανότητες $30\%, 10\%, 60\%$. Θα υπολογίσουμε τη μέση τιμή και τη διασπορά του κέρδους σε μια παρτίδα. Εφαρμόζοντας τους ορισμούς,

$$E(X) = (-20) \times \frac{30}{100} + 0 \times \frac{10}{100} + 10 \times \frac{60}{100} = -6 + 0 + 6 = 0.$$

και

$$\text{VAR}(X) = E(X^2) - (E(X))^2 = (-20)^2 \times \frac{30}{100} + 0^2 \times \frac{10}{100} + 10^2 \times \frac{60}{100} = 180.$$

Πόση είναι η μέση τιμή και η διασπορά των κερδών αν ο παίκτης παίξει N παρτίδες; Επί του παρόντος, δεν έχουμε τη θεωρία για να απαντήσουμε το άνω ερώτημα εύκολα. Σε επόμενα κεφάλαια θα δείξουμε πως τα ερωτήματα αυτού του τύπου μπορούν, υπό προϋποθέσεις, να απαντηθούν πολύ εύκολα.

Παράδειγμα 4.12. *(Υπολογισμός παραμέτρων)* Μια Τ.Μ. X παίρνει τις τιμές $0, 1, 2$ και k. Η μάζα της είναι $p(0) = 1/6$, $p(1) = 1/5$, $p(2) = 1/3$ και $p(k) = a$. Θα απαντήσουμε στα ακόλουθα ερωτήματα:

1. Υπολογίστε την τιμή του a.
2. Αν η μέση τιμή του X είναι $8/3$, υπολογίστε την τιμή του k.
3. Υπολογίστε την διασπορά του X με βάση τις πιο πάνω τιμές.

Έχουμε, κατά περίπτωση:

1. Θα πρέπει

$$p(0) + p(1) + p(2) + p(k) = 1 \Rightarrow \frac{1}{6} + \frac{1}{5} + \frac{1}{3} + a = 1 \Rightarrow a = \frac{3}{10}.$$

2.

$$E(X) = 0 \times p(0) + 1 \times p(1) + 2 \times p(2) + k \times p(k) = 0 + \frac{1}{5} + \frac{2}{3} + \frac{3k}{10} = \frac{26 + 9k}{30}.$$

Όμως δίνεται ότι $E(X) = 8/3$, άρα έχουμε

$$\frac{26 + 9k}{30} = \frac{8}{3} \Rightarrow k = 6.$$

3.

$$\text{VAR}(X) = \frac{1}{6}\left(0 - \frac{8}{3}\right)^2 + \frac{1}{5}\left(1 - \frac{8}{3}\right)^2 + \frac{1}{3}\left(2 - \frac{8}{3}\right)^2 + \frac{3}{10}\left(6 - \frac{8}{3}\right)^2 = \frac{47}{9}.$$

Παράδειγμα 4.13. *(Ομοιόμορφη διακριτή Τ.Μ.)* Έστω η Τ.Μ. X που παίρνει τιμές από το σύνολο $\{1, 2, \dots, n\}$ με ίσες πιθανότητες. Αυτή η Τ.Μ. είναι γνωστή ως ομοιόμορφη. Θα υπολογίσουμε καταρχήν την μέση τιμή και τη διασπορά της.

Η X έχει τη μάζα πιθανότητας $p_X(i) = \frac{1}{n}$, $i = 1, 2, \dots, n$. Από τον ορισμό της μέσης τιμής,

$$E(X) = \sum_{i=1}^{n} i\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} i = \frac{n+1}{2}.$$

Στην τελευταία ισότητα χρησιμοποιήσαμε τη γνωστή ταυτότητα

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

Επιπλέον,

$$E(X^2) = \sum_{i=1}^{n} i^2\frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} i^2 = \frac{(n+1)(2n+1)}{6}.$$

Στην τελευταία ισότητα χρησιμοποιήσαμε τη γνωστή ταυτότητα

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

Συνεπώς:

$$\text{VAR}(X) = E(X^2) - (E(X))^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}.$$

Έστω τώρα η Τ.Μ. $Y = aX + b$. Η μέση τιμή της και η διασπορά της ισούνται με

$$\begin{aligned} E(Y) &= E(a+bX) = a + bE(X) = a + b\frac{n+1}{2}, \\ \text{VAR}(Y) &= \text{VAR}(a+bX) = b^2\text{VAR}(X) = b^2\frac{n^2-1}{12}. \end{aligned}$$

Στα άνω χρησιμοποιήσαμε τα Λήμματα 4.4 και 4.5.

Παράδειγμα 4.14. *(Κανονικοποίηση)* Αν η X είναι Τ.Μ. με μέση τιμή $E(X) = \mu$ και διασπορά $\text{VAR}(X) = \sigma^2$, τότε η Τ.Μ. $Y = \frac{X-\mu}{\sigma}$ έχει μέση τιμή $E(Y) = 0$ και διασπορά $\text{VAR}(Y) = 1$. Πράγματι,

$$\begin{aligned} E(Y) &= E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(X) - \mu\right] = 0, \\ \text{VAR}(Y) &= \text{VAR}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\text{VAR}(X) = 1. \end{aligned}$$

Κεφάλαιο 5

Συνήθεις Περιπτώσεις Διακριτών Τυχαίων Μεταβλητών

Στο προηγούμενο κεφάλαιο είδαμε πως η έννοια της τυχαίας μεταβλητής (Τ.Μ.), δηλαδή μιας «τυχαίας ποσότητας» X που προσδιορίζεται από το σύνολο τιμών της S και την μάζα της $p(\cdot)$, μας επιτρέπει να περιγράφουμε πολλά από τα προβλήματα που μας ενδιαφέρουν με σύντομο και σαφή τρόπο, χωρίς να αναγκαζόμαστε κάθε φορά να ορίζουμε σχολαστικά τον πλήρη δειγματικό χώρο Ω και το αντίστοιχο μέτρο πιθανότητας P. Το επόμενο βήμα σ′ αυτήν τη διαδικασία της πιο «αφαιρετικής» περιγραφής είναι η καταγραφή κάποιων σημαντικών κατηγοριών Τ.Μ., οι οποίες είναι ιδιαίτερα χρήσιμες στην πράξη και εμφανίζονται συχνά σε βασικά προβλήματα των πιθανοτήτων.

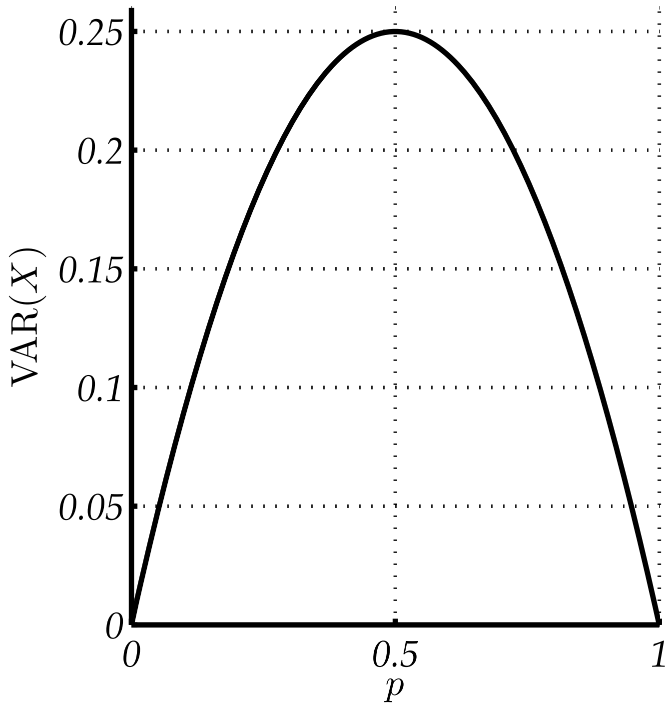
Σ′ αυτό το κεφάλαιο θα εντοπίσουμε πέντε τέτοιες κατηγορίες Τ.Μ., και θα δείξουμε με παραδείγματα κάποιες από τις συνήθεις περιπτώσεις όπου χρησιμοποιούνται. Επιπλέον θα αποδείξουμε ορισμένες ιδιότητές τους (π.χ., θα υπολογίσουμε τη μέση τιμή και τη διασπορά τους), έτσι ώστε να μπορούμε να τις χρησιμοποιούμε χωρίς να απαιτείται η επανάληψη των ίδιων υπολογισμών σε κάθε νέο πρόβλημα.

5.1 Κατανομή Bernoulli

Η πιο απλή μη τετριμμένη Τ.Μ. είναι εκείνη που παίρνει μόνο δύο τιμές, οι οποίες, στην απλούστερη περίπτωση, είναι το 0 και το 1:

Ορισμός 5.1. (Κατανομή Bernoulli) *Μια διακριτή Τ.Μ. X λέμε πως ακολουθεί την κατανομή Bernoulli με παράμετρο p, για κάποιο $p \in (0,1)$, αν έχει σύνολο τιμών το $S_X = \{0,1\}$ και μάζα p_X με τιμές $p_X(1) = p$ και $p_X(0) = 1-p$. Για συντομία, γράφουμε $X \sim \mathrm{Bern}(p)$.*

Παρατήρηση: Η κατανομή Bernoulli χρησιμοποιείται πολύ συχνά όταν έχουμε ένα πείραμα που μπορεί να πάρει 2 δυνατά αποτελέσματα, για παράδειγμα:



Σχήμα 5.1: Η διασπορά της κατανομής Bernoulli συναρτήσει της παραμέτρου p.

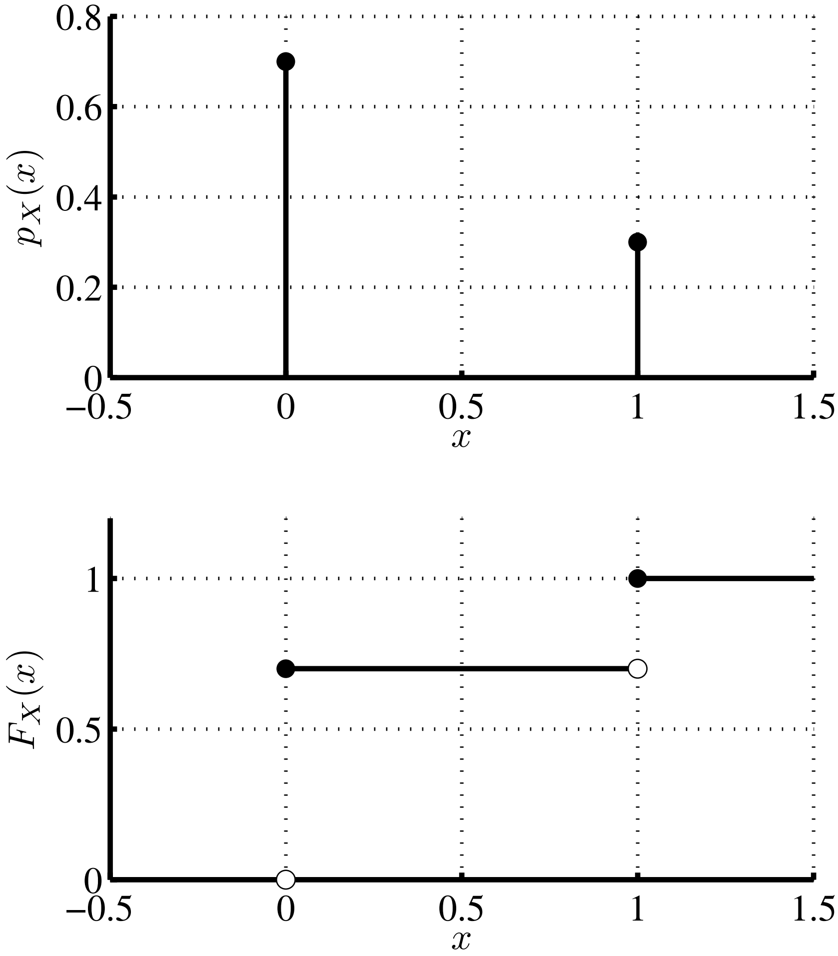
1. Ρίχνουμε ένα κέρμα, με δυνατά αποτελέσματα κορώνα/γράμματα.
2. Κάνουμε ένα παιδί, με δυνατά αποτελέσματα κορίτσι/αγόρι.
3. Εκτελούμε ένα πέναλτι και εξετάζουμε αν μπήκε γκολ ή όχι.
4. Κάνουμε ένα πείραμα (που ενδεχομένως έχει πολλά δυνατά αποτελέσματα) και εξετάζουμε αν προέκυψε κάποιο ενδεχόμενο E ή το E'.

Σε αυτές τις περιπτώσεις, κωδικοποιούμε το ένα αποτέλεσμα με 0, και το αντιστοιχούμε σε αποτυχία, και το άλλο με 1, και το αντιστοιχούμε σε επιτυχία.

Λήμμα 5.1. (Ιδιότητες κατανομής Bernoulli) *Έστω* $X \sim \text{Bern}(p)$.

1. $E(X) = p$.
2. $\text{VAR}(X) = p(1-p)$.
3. *Η κατανομή της* X *ισούται με*

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1-p, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$



Σχήμα 5.2: Η μάζα και η κατανομή Bernoulli για $p = \frac{1}{3}$.

Απόδειξη. 1. Εξ ορισμού,

$$E(X) = 0 \times (1 - p) + 1 \times p = p.$$

2. Έχουμε

$$E(X^2) = 0^2 \times (1 - p) + 1^2 \times p = p$$
$$\Rightarrow \mathrm{VAR}(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p).$$

   Η διασπορά έχει σχεδιαστεί στο Σχήμα 5.1. Παρατηρήστε πως είναι μέγιστη για $p = \frac{1}{2}$ και μηδενική για $p = 0$ και $p = 1$.

3. Προφανές, εξετάζοντας την κάθε περίπτωση χωριστά. Η μάζα και η κατανομή έχουν σχεδιαστεί στο Σχήμα 5.2.

□

5.2 Διωνυμική Κατανομή

Παράδειγμα 5.1. *(4 ζάρια)* Έστω πως ρίχνουμε τέσσερα δίκαια ζάρια, και οι ρίψεις είναι ανεξάρτητες. Ποιες είναι οι πιθανότητες να φέρουμε i φορές 6, για $i = 0, 1, 2, 3$;

Έστω X_1, X_2, X_3, X_4 τα αποτελέσματα των ζαριών, και X το πλήθος των ζαριών που ήρθαν 6.

Παρατηρούμε πως

$$\begin{aligned} P(X=0) &= P(X_1, X_2, X_3, X_4 \neq 6) \\ &= P(X_1 \neq 6)P(X_2 \neq 6)P(X_3 \neq 6)P(X_4 \neq 6) = \left(\frac{5}{6}\right)^4 . \end{aligned}$$

Η δεύτερη ισότητα προέκυψε γιατί υποθέσαμε ότι οι 4 ζαριές είναι ανεξάρτητες μεταξύ τους, δηλαδή ενδεχόμενα που αναφέρονται σε διαφορετικά ζάρια είναι ανεξάρτητα. Επιπλέον,

$$\begin{aligned} P(X=1) &= P(X_1 = 6, X_2, X_3, X_4 \neq 6) + P(X_2 = 6, X_1, X_3, X_4 \neq 6) \\ &\quad + P(X_3 = 6, X_1, X_2, X_4 \neq 0) + P(X_4 = 6, X_1, X_2, X_3 \neq 6) \\ &= P(X_1 = 6)P(X_2 \neq 6)P(X_3 \neq 6)P(X_4 \neq 6) \\ &\quad + P(X_1 \neq 6)P(X_2 = 6)P(X_3 \neq 6)P(X_4 \neq 6) \\ &\quad + P(X_1 \neq 6)P(X_2 \neq 6)P(X_3 = 6)P(X_4 \neq 6) \\ &\quad + P(X_1 \neq 6)P(X_2 \neq 6)P(X_3 \neq 6)P(X_4 = 6) \\ &= \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + \frac{1}{6} \times \left(\frac{5}{6}\right)^3 \\ &= 4 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 . \end{aligned}$$

Για την περίπτωση των δύο επιτυχιών, γράφοντας πιο συνοπτικά έχουμε

$$\begin{aligned} P(X=2) &= P(\not 6\ \not 6\ 6\ 6) + P(\not 6\ 6\ \not 6\ 6) + P(\not 6\ 6\ 6\ \not 6) \\ &\quad + P(6\ \not 6\ 6\ \not 6) + P(6\ 6\ \not 6\ \not 6) + P(6\ \not 6\ \not 6\ 6) = 6\left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right)^2 . \end{aligned}$$

Οι περιπτώσεις των 3 και 4 επιτυχιών προκύπτουν ανάλογα. Παρατηρήστε ότι το αποτέλεσμα είναι πάντα το άθροισμα ίσων όρων, καθένας εκ των οποίων εκφράζει την πιθανότητα να προκύψουν i επιτυχίες με ένα συγκεκριμένο τρόπο. Η συγκεκριμένη μεθοδολογία γενικεύεται στο ακόλουθο θεώρημα.

Λήμμα 5.2. (*N* ανεξάρτητα πειράματα) *Έστω πως διεξάγουμε N πειράματα, καθένα εκ των οποίων μπορεί να καταλήξει σε επιτυχία ή αποτυχία με πιθανότητα p, ανεξάρτητα από την έκβαση των υπόλοιπων πειραμάτων. Έστω X το πλήθος των επιτυχιών. Η πιθανότητα να έχουμε ακριβώς x επιτυχίες, όπου $x = 0, 1, \ldots, N$, ισούται με*

$$P(X = x) = \binom{N}{x} p^x (1-p)^{N-x}.$$

Απόδειξη. Έστω η Τ.Μ. Bernoulli Y_i που εκφράζει το αποτέλεσμα του πειράματος i. Έστω πως $Y_i = 1$ όταν το i πείραμα είναι επιτυχία, και $Y_i = 0$ όταν είναι αποτυχία. Προφανώς, $P(Y_i = 1) = p$, $P(Y_i = 0) = 1 - p$.

Έστω ω το αποτέλεσμα να προκύψει μια συγκεκριμένη ακολουθία από x επιτυχίες και $N - x$ αποτυχίες, που μπορεί να περιγραφεί ως $Y_i = y_i$, $i = 1, 2, \ldots, N$, όπου x από τα y_i είναι 1, και τα υπόλοιπα είναι 0. Η πιθανότητα αυτού του αποτελέσματος ισούται με

$$\begin{aligned} P(\omega) &= P(Y_1 = y_1,\ Y_2 = y_2, \ldots, Y_N = y_N) \\ &= P(Y_1 = y_1)P(Y_2 = y_2) \ldots P(Y_N = y_N) \\ &= p^x (1-p)^{N-x}. \end{aligned}$$

Η δεύτερη ισότητα προκύπτει λόγω ανεξαρτησίας των πειραμάτων. Η τρίτη προκύπτει γιατί στο άνω γινόμενο έχουμε x παράγοντες που αντιστοιχούν σε επιτυχίες, καθέναν ίσο με p, και $N - x$ παράγοντες που αντιστοιχούν σε αποτυχίες, καθέναν ίσο με $1 - p$.

Με πόσους όμως τρόπους μπορώ να επιλέξω τη σειρά των x επιτυχιών; Το πρόβλημα είναι ισοδύναμο με την επιλογή x στοιχείων από N, χωρίς επανάθεση και χωρίς να μας ενδιαφέρει η σειρά. Άρα, υπάρχουν $\binom{N}{x}$ αποτελέσματα που οδηγούν σε x επιτυχίες, όλα με πιθανότητα $p^x(1-p)^{N-x}$. Άρα, η πιθανότητα για x επιτυχίες είναι ως άνω. □

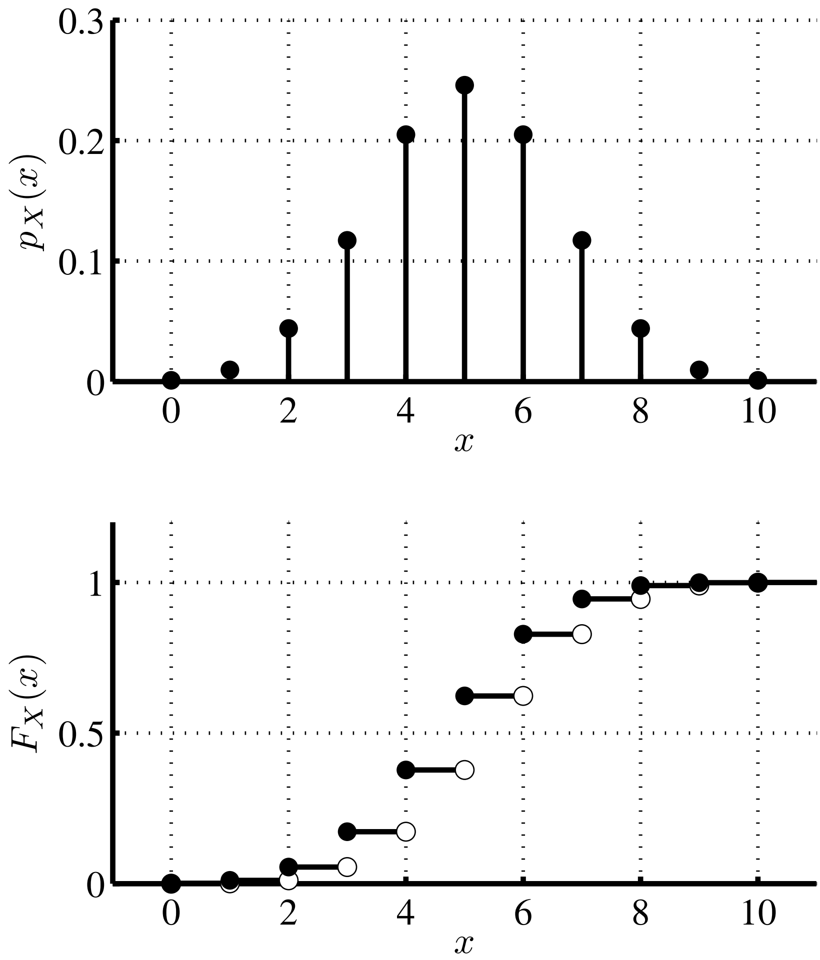
Παρατήρηση: Η περίπτωση όπου έχουμε N ανεξάρτητα πειράματα, καθένα εκ των οποίων είναι ανεξάρτητο από τα υπόλοιπα, και όλα έχουν την ίδια πιθανότητα επιτυχίας p προκύπτει πολύ συχνά. Για παράδειγμα:

1. Κάνουμε N παιδιά και μετράμε πόσα ήταν κορίτσια.

2. Εκτελούμε N πέναλτι και μετράμε πόσα μπήκαν.

3. Πυροβολούμε προς ένα στόχο N φορές, και μετράμε τις φορές που τον πετύχαμε.

Ακριβώς λόγω της συχνότητας που εμφανίζεται αυτή η περίπτωση, η μελέτη του αριθμού των επιτυχιών παρουσιάζει ιδιαίτερο ενδιαφέρον.



Σχήμα 5.3: Η μάζα και η κατανομή μιας διωνυμικής Τ.Μ. με παραμέτρους $N = 10$, $p = 0.5$.

Ορισμός 5.2. (Διωνυμική κατανομή) *Μια διακριτή τυχαία μεταβλητή X ακολουθεί την διωνυμική κατανομή με παραμέτρους N και p, για κάποια $N \geq 1$ και $p \in (0, 1)$, αν έχει σύνολο τιμών το $S_X = \{0, 1, \ldots, N\}$ και μάζα:*

$$p_X(x) = \binom{N}{x} p^x (1-p)^{N-x}, \quad x \in S_X = \{0, 1, \ldots, N\}.$$

Για συντομία, γράφουμε $X \sim$ Διων(N, p).

Λήμμα 5.3. (Ιδιότητες διωνυμικής κατανομής) *Έστω X διωνυμική Τ.Μ. με παραμέτρους N, p.*

1. $E(X) = Np$.
2. $\mathrm{VAR}(X) = Np(1-p)$.

Απόδειξη. Για τον υπολογισμό της μέσης τιμής, παρατηρούμε πως

$$\begin{aligned} E(X) &= \sum_{x=0}^{N} x p_X(x) = \sum_{x=0}^{N} x \binom{N}{x} p^x (1-p)^{N-x} \\ &= \sum_{x=1}^{N} \frac{N!}{(N-x)!(x-1)!} p^x (1-p)^{N-x} \\ &= Np \sum_{x=1}^{N} \frac{(N-1)!}{(N-x)!(x-1)!} p^{x-1} (1-p)^{N-x} \\ &= Np \sum_{j=0}^{N-1} \binom{N-1}{j} p^j (1-p)^{N-1-j} = Np[p+(1-p)]^{N-1} = Np. \end{aligned}$$

Στην πέμπτη ισότητα κάναμε την αντικατάσταση $j = x-1$, ενώ στην τελευταία χρησιμοποιήσαμε το διωνυμικό θεώρημα:

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

Για τον υπολογισμό της διασποράς έχουμε καταρχήν:

$$\begin{aligned} E(X(X-1)) &= \sum_{x=0}^{N} x(x-1) \binom{N}{x} p^x (1-p)^{N-x} \\ &= \sum_{x=2}^{N} \frac{N!}{(N-x)!(x-2)!} p^x (1-p)^{N-x} \\ &= N(N-1)p^2 \sum_{x=2}^{N} \frac{(N-2)!}{(N-x)!(x-2)!} p^{x-2} (1-p)^{N-x} \\ &= N(N-1)p^2 \sum_{j=0}^{N-2} \frac{(N-2)!}{(N-j-2)!j!} p^j (1-p)^{N-j-2} \\ &= N(N-1)p^2 [p+(1-p)]^{N-2} = N(N-1)p^2. \end{aligned}$$

Η τέταρτη ισότητα προκύπτει με την αντικατάσταση $j = x-2$, ενώ η πέμπτη και πάλι με χρήση του διωνυμικού θεωρήματος. Με χρήση αυτής της ισότητας, έχουμε

$$E(X^2) = E(X(X-1)) + E(X) = N(N-1)p^2 + Np,$$

και τελικά

$$\mathrm{VAR}(X) = E(X^2) - (E(X))^2 = N(N-1)p^2 + Np - N^2p^2 = Np(1-p).$$

□

Παρατηρήσεις

1. Όπως φαίνεται και από την προηγούμενη συζήτηση, μια διωνυμική Τ.Μ. με παραμέτρους N, p μπορεί να εκφρασθεί όπως το άθροισμα N Τ.Μ. Bernoulli οι οποίες έχουν πιθανότητα επιτυχίας p, και αντιστοιχούν σε ανεξάρτητα μεταξύ τους πειράματα.

2. Δεν υπάρχει κάποιος απλός τύπος που να δίνει την διωνυμική συνάρτηση κατανομής πιθανότητας. Ένα παράδειγμα έχει σχεδιαστεί, μαζί με την αντίστοιχη μάζα, στο Σχήμα 5.3.

Παράδειγμα 5.2. *(Κατανομή μηνυμάτων spam)* Από τα μηνύματα email που λαμβάνει ένας χρήστης, το 35% είναι διαφημιστικά μηνύματα spam. Υποθέτουμε ότι το αν κάποιο μήνυμα είναι spam ή όχι είναι ανεξάρτητο από μήνυμα σε μήνυμα. Έστω ότι ο χρήστης λαμβάνει συνολικά 25 email σε μία μέρα, και έστω X το πλήθος των spam ανάμεσα σ′ αυτά.

Η κατανομή της X είναι η διωνυμική, με παραμέτρους $N = 25$ και $p = 0.35$, αν θεωρήσουμε ότι η λήψη μηνύματος είναι πείραμα που επιτυγχάνει όταν το μήνυμα είναι spam. Έτσι, για παράδειγμα, η πιθανότητα ο χρήστης να λάβει σε μια μέρα 8, 9 ή 10 μηνύματα ισούται με

$$\begin{aligned}
P(X = 8, 9, 10) &= p_X(8) + p_X(9) + p_X(10) \\
&= \binom{25}{8} 0.35^8 0.65^{17} + \binom{25}{9} 0.35^9 0.65^{16} + \binom{25}{10} 0.35^{10} 0.65^{15} \\
&\simeq 0.4651.
\end{aligned}$$

Παράδειγμα 5.3. *(Overbooking)* Ένα αεροπλάνο έχει 200 θέσεις, και έχουν γίνει κρατήσεις από 210 επιβάτες. Κάθε ένας από αυτούς θα έρθει στο αεροδρόμιο με πιθανότητα 93%, ανεξάρτητα από τους υπόλοιπους. Ποια είναι η πιθανότητα κάποιοι επιβάτες που θα φθάσουν στο αεροδρόμιο να είναι υπεράριθμοι;

Παρατηρούμε πως το πλήθος X των επιβατών που θα προσέλθουν ακολουθεί διωνυμική κατανομή με παραμέτρους το πλήθος των ανεξάρτητων πειραμάτων $N = 210$ και πιθανότητα επιτυχίας $p = 0.93$. Θα έχουμε πρόβλημα υπεράριθμων επιβατών αν προσέλθουν 201 με 210 επιβάτες, δηλαδή αν $X \geq 201$. Η πιθανότητα αυτή μπορεί εύκολα να υπολογιστεί από τον τύπο της μάζας της διωνυμικής κατανομής:

$$\begin{aligned}
P(X \geq 201) &= P(X = 201) + P(X = 201) + \cdots + P(X = 210) \\
&= \binom{210}{201} 0.93^{201} 0.07^9 + \binom{210}{202} 0.93^{202} 0.07^8 \\
&\quad + \cdots + \binom{210}{210} 0.93^{210} 0.07^0 \\
&\simeq 0.0728.
\end{aligned}$$

5.3 Χρήσιμες Ταυτότητες

Πριν ορίσουμε την γεωμετρική κατανομή, θα αποδείξουμε μερικές σχέσεις που εμφανίζονται συχνά σε σχετικούς της υπολογισμούς. Πιο συγκεκριμένα, θα αποδείξουμε το ακόλουθο λήμμα:

Λήμμα 5.4. (Η γεωμετρική και συναφείς σειρές) *Ισχύουν οι ακόλουθες σχέσεις, για κάθε* $y \in \mathbb{R}$ *με* $|y| < 1$:

$$\sum_{k=0}^{\infty} y^k = \frac{1}{1-y}, \tag{5.1}$$

$$\sum_{k=1}^{\infty} k y^k = \frac{y}{(1-y)^2}, \tag{5.2}$$

$$\sum_{k=1}^{\infty} k^2 y^k = \frac{y(1+y)}{(1-y)^3}. \tag{5.3}$$

Απόδειξη. Ξεκινάμε από την απλή παρατήρηση ότι, για κάθε $y \in \mathbb{R}$, και κάθε $n \geq 1$,

$$(1 + y + y^2 + \cdots + y^n)(1-y) = 1 - y^{n+1}.$$

(Αν αυτή η σχέση δεν σας είναι προφανής, αποδείξτε την με επαγωγή.) Υποθέτοντας ότι $y \neq 1$ και διαιρώντας και τα δύο μέρη με το $(1-y)$, έχουμε τον γνωστό τύπο για το άθροισμα των όρων μιας γεωμετρικής προόδου:

$$\sum_{k=0}^{n} y^k = \frac{1-y^{n+1}}{1-y}, \quad \text{για κάθε } y \neq 1.$$

Αν περιοριστούμε τώρα σε τιμές του y με $|y| < 1$, τότε προφανώς το $y^{n+1} \to 0$ καθώς το $n \to \infty$, άρα, περνώντας στο όριο έχουμε:

$$\sum_{k=0}^{\infty} y^k \triangleq \lim_{n\to\infty} \sum_{k=0}^{n} y^k = \lim_{n\to\infty} \frac{1-y^{n+1}}{1-y} = \frac{1 - \lim_{n\to\infty} y^{n+1}}{1-y} = \frac{1}{1-y},$$

και αποδείξαμε την (5.1).

Ορίζουμε τώρα τη συνάρτηση

$$g(y) \triangleq \sum_{k=0}^{\infty} y^k = \frac{1}{1-y}, \quad |y| < 1.$$

Έχουμε δύο διαφορετικές εκφράσεις για την $g(y)$, μια ως σειρά και μια «σε κλειστή μορφή». Άρα μπορούμε να υπολογίσουμε την παράγωγο $g'(y)$ με δύο τρόπους: Αφού η παράγωγος ενός αθροίσματος ισούται με το άθροισμα των παραγώγων,

$$g'(y) = \frac{d}{dy}\left(\sum_{k=0}^{\infty} y^k\right) = \sum_{k=0}^{\infty} \frac{d}{dy}(y^k) = \sum_{k=0}^{\infty} k y^{k-1} = \sum_{k=1}^{\infty} k y^{k-1},$$

αλλά και

$$g'(y) = \frac{d}{dy}\Big(\frac{1}{1-y}\Big) = \frac{1}{(1-y)^2}.$$

Εξισώνοντας τις δύο πιο πάνω εκφράσεις και πολλαπλασιάζοντας και τα δύο μέρη με το y, προκύπτει τελικά η (5.2).

Για να αποδείξουμε την (5.3), ορίζουμε τη συνάρτηση

$$h(y) \triangleq \sum_{k=1}^{\infty} k y^k = \frac{y}{(1-y)^2}, \quad |y| < 1.$$

Επαναλαμβάνοντας την ίδια διαδικασία μπορούμε να υπολογίσουμε την παράγωγο της νέας συνάρτησης $h(y)$ με δύο τρόπους, ως

$$\begin{aligned} h'(y) &= \frac{d}{dy}\left(\sum_{k=1}^{\infty} k y^k\right) = \sum_{k=1}^{\infty} \frac{d}{dy}(k y^k) = \sum_{k=1}^{\infty} k^2 y^{k-1}, \\ h'(y) &= \frac{d}{dy}\Big(\frac{y}{(1-y)^2}\Big) = \frac{(1-y)^2 + 2y(1-y)}{(1-y)^4} = \frac{1+y}{(1-y)^3}. \end{aligned}$$

Εξισώνοντας όπως πριν τις δύο εκφράσεις για την παράγωγο και πολλαπλασιάζοντας και τα δύο μέρη με το y, προκύπτει και η (5.3). □

Παρατήρηση: Μπορείτε να βρείτε στην άνω απόδειξη κάποια σημεία που δεν ήμασταν εντελώς αυστηροί; (Περισσότερα σε προχωρημένα μαθήματα μαθηματικής ανάλυσης.)

5.4 Γεωμετρική Κατανομή

Παράδειγμα 5.4. *(Ακολουθία από ζαριές)* Έστω πως ρίχνουμε ένα δίκαιο ζάρι διαδοχικά μέχρι να φέρουμε για πρώτη φορά 6. Το αποτέλεσμα μιας ρίψεως δεν επηρεάζει το αποτέλεσμα των επόμενων. Έστω E_x, όπου $x = 1, 2, \ldots$, το ενδεχόμενο να φέρουμε 6 για πρώτη φορά στην προσπάθεια x. Θα υπολογίσουμε την πιθανότητα $P(E_x)$.

Παρατηρήστε πως:

$$P(E_1) = \frac{1}{6}.$$

Επιπλέον,

$$P(E_2) = P(E_1'E_2) = P(E_1')P(E_2|E_1') = \frac{5}{6} \times \frac{1}{6}.$$

Με παρόμοιο τρόπο,

$$P(E_3) = P(E_1'E_2'E_3) = P(E_1')P(E_2'|E_1')P(E_3|E_1'E_2') = \left(\frac{5}{6}\right)^2 \frac{1}{6},$$

και πιο γενικά

$$P(E_x) = P(E_1'E_2' \ldots E_{x-1}'E_x) = \\ P(E_1')P(E_2'|E_1') \ldots P(E_x|E_1'E_2' \ldots E_{x-1}') = \left(\frac{5}{6}\right)^{x-1} \frac{1}{6}.$$

Παρατηρήσεις

1. Το άνω μαθηματικό πρόβλημα εμφανίζεται πολύ συχνά, σε πληθώρα εφαρμογών. Ουσιαστικά επαναλαμβάνουμε πολλές φορές ένα πείραμα που μπορεί να προκύψει αποτυχία ή επιτυχία, μέχρι να έχουμε επιτυχία, και μας ενδιαφέρει να μελετήσουμε σε ποια προσπάθεια θα έχουμε την επιτυχία. Για παράδειγμα:

   (α΄) Κάνουμε παιδιά μέχρι να κάνουμε το πρώτο κορίτσι.

   (β΄) Δίνουμε ένα μάθημα μέχρι να το περάσουμε.

   (γ΄) Στέλνουμε ένα πακέτο σε ένα server μέχρι να το λάβει επιτυχώς.

2. Πιο γενικά, έχουμε το ακόλουθο λήμμα.

Λήμμα 5.5. (Ακολουθία από πειράματα μέχρι την πρώτη επιτυχία) *Έστω πως εκτελούμε μια ακολουθία πειραμάτων που είναι ανεξάρτητα μεταξύ τους, με το καθένα να είναι επιτυχημένο με πιθανότητα p. Έστω E_x το ενδεχόμενο το πρώτο επιτυχημένο πείραμα να είναι στην προσπάθεια x, όπου $x = 1, 2, \ldots$ Τότε*

$$P(E_x) = (1-p)^{x-1}p, \quad x = 1, 2, \ldots \tag{5.4}$$

Απόδειξη. Παρατηρήστε πως

$$\begin{aligned} P(E_x) &= P(E_1'E_2'\dots E_{x-1}'E_x) \\ &= P(E_1')P(E_2'|E_1')\dots P(E_x|E_1'E_2'\dots E_{x-1}') = (1-p)^{x-1}p. \end{aligned}$$

□

Παρατηρήσεις

1. Μπορούμε να παρατηρήσουμε πως

$$\sum_{x=1}^{\infty}(1-p)^{x-1}p = p \times \sum_{i=0}^{\infty}(1-p)^i = p \times \frac{1}{1-(1-p)} = 1.$$

   Στη δεύτερη ισότητα θέσαμε $i = x-1$. Στην τρίτη ισότητα, χρησιμοποιήσαμε την (5.1) κάνοντας την αντικατάσταση $y \to (1-p)$, $k \to i$. Άρα, θα γίνει πεπερασμένος αριθμός πειραμάτων με πιθανότητα 1.

2. Παρατηρήστε πως το άνω λήμμα ισχύει είτε σταματάμε στο πρώτο επιτυχημένο πείραμα, όπως στα άνω παραδείγματα, είτε συνεχίζουμε τα πειράματα επ' άπειρον, είτε τα σταματάμε οποιαδήποτε στιγμή μετά την πρώτη επιτυχία.

Ορισμός 5.3. (Γεωμετρική κατανομή) *Μια διακριτή τυχαία μεταβλητή X λέμε πως έχει γεωμετρική κατανομή με παράμετρο p, για κάποιο $p \in (0,1)$, αν έχει σύνολο τιμών το $S_X = \{1, 2, \dots\}$ και μάζα*

$$p_X(x) = (1-p)^{x-1}p, \quad x \in S_X = \{1, 2, \dots\}.$$

Για συντομία, γράφουμε $X \sim$ Γεωμ(p).

Λήμμα 5.6. (Ιδιότητες γεωμετρικής κατανομής) *Έστω $X \sim$ Γεωμ(p).*

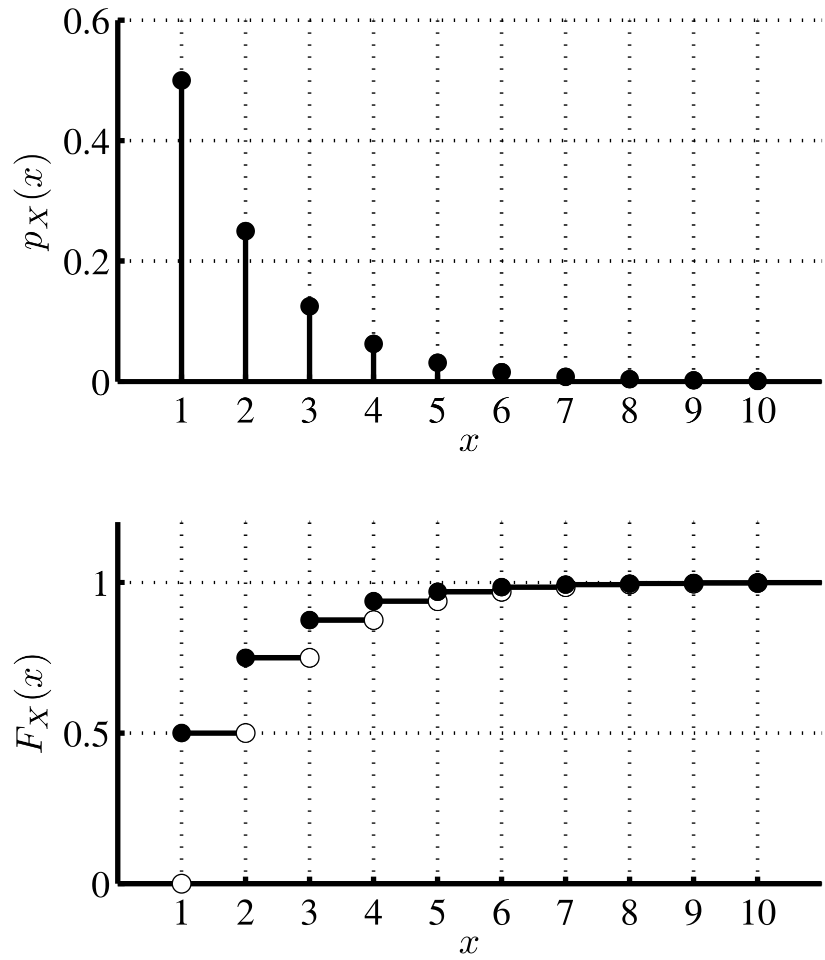
1. *Για κάθε ακέραιο $m \geq 0$, ισχύει ότι $P(X > m) = (1-p)^m$.*

2. *Η συνάρτηση κατανομής πιθανότητας της X ισούται με*

$$F_X(X) = \begin{cases} 1-(1-p)^{\lfloor x \rfloor}, & x \geq 1, \\ 0, & x < 1. \end{cases}$$

3. *Η μέση τιμή της X ισούται με $E(X) = 1/p$.*

4. *Η διασπορά της X ισούται με $\mathrm{VAR}(X) = (1-p)/p^2$.*



Σχήμα 5.4: Η μάζα και η κατανομή μιας γεωμετρικής Τ.Μ. με παραμέτρους $p = 0.5$.

5. *Για κάθε ζευγάρι ακεραίων $m \geq 1$ και $n \geq 0$:*

$$\begin{aligned} P(X = m + n \,|\, X > n) &= P(X = m), \\ P(X \geq m + n \,|\, X > n) &= P(X \geq m), \\ P(X > m + n \,|\, X > n) &= P(X > m). \end{aligned}$$

Απόδειξη. 1. Η περίπτωση $m = 0$ είναι προφανής. Αν τώρα $m > 0$, παρατηρήστε πως θα έχουμε $X > m$ αν και μόνο αν οι πρώτες m προσπάθειες είναι όλες αποτυχίες. Αυτό γίνεται με πιθανότητα $(1 - p)^m$.

2. Παίρνουμε περιπτώσεις: όταν $x < 1$, προφανώς $F_X(x) = P(X \leq x) = 0$. Αν πάλι το x είναι θετικός ακέραιος, τότε

$$F_X(x) = P(X \leq x) = 1 - P(X > x) = 1 - (1 - p)^x = 1 - (1 - p)^{\lfloor x \rfloor},$$

από το προηγούμενο σκέλος. Τέλος, αν το $x > 1$ αλλά όχι ακέραιο, τότε απλώς παρατηρούμε πως, αφού το X είναι ακέραιο, θα έχουμε $X \leq x$ αν και μόνο αν $X \leq \lfloor x \rfloor$, άρα και πάλι μπορούμε να χρησιμοποιήσουμε την άνω σχέση. Η κατανομή και η μάζα έχουν σχεδιαστεί στο Σχήμα 5.4 για την περίπτωση $p = 0.5$.

3. Παρατηρούμε πως

$$\begin{aligned} E(X) &= \sum_{x=1}^{\infty} x p_X(x) = \sum_{x=0}^{\infty} x(1-p)^{x-1} p = \frac{p}{1-p} \sum_{x=1}^{\infty} x(1-p)^x \\ &= \frac{p}{1-p} \times \frac{1-p}{[1-(1-p)]^2} = \frac{1}{p}, \end{aligned}$$

όπου στην δεύτερη από το τέλος ισότητα εφαρμόσαμε την (5.2) με $k \to x$, $y \to 1-p$.

4. Παρομοίως,

$$\begin{aligned} E(X^2) &= \sum_{x=1}^{\infty} x^2 p_X(x) = \sum_{x=0}^{\infty} x^2(1-p)^{x-1} p = \frac{p}{1-p} \sum_{x=1}^{\infty} x^2(1-p)^x \\ &= \frac{p}{1-p} \times \frac{(1-p)(1+1-p)}{[1-(1-p)]^3} = \frac{2-p}{p^2}, \\ \mathrm{VAR}(X) &= E(X^2) - [E(X)]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}. \end{aligned}$$

Στην πρώτη ισότητα της δεύτερης γραμμής εφαρμόσαμε την (5.3) με $k \to x$, $y \to 1-p$.

5. Θα αποδείξουμε μόνο την πρώτη ισότητα, καθώς οι αποδείξεις των άλλων είναι παρόμοιες. Απλώς παρατηρούμε πως

$$\begin{aligned} P(X = m+n | X > n) &= \frac{P(X = m+n, X > n)}{P(X > n)} = \frac{P(X = m+n)}{P(X > n)} \\ &= \frac{p(1-p)^{m+n-1}}{(1-p)^n} = p(1-p)^{m-1} = P(X = m). \end{aligned}$$

Δηλαδή από την πρώτη ισότητα προκύπτει ότι η πιθανότητα $P(X = m+n \,|\, X > n)$ είναι ανεξάρτητη του n και ίση με $P(X = m)$. Παρόμοια εξηγούνται και οι άλλες ισότητες.

Διαισθητικά, αν γνωρίζουμε ότι είχαμε στην αρχή n αποτυχίες, τότε η περιγραφή που πειράματος από εκεί και πέρα είναι ίδια με την αρχική περιγραφή του πειράματος. Η ιδιότητα αυτή καλείται *έλλειψη μνήμης*.

□

Παράδειγμα 5.5. Σε ένα iPod shuffle υπάρχουν 100 τραγούδια, που παίζονται με τυχαία σειρά, και κάθε τραγούδι επιλέγεται με πιθανότητα 1/100, και ανεξάρτητα από τις προηγούμενες επιλογές. Από τα 100 τραγούδια, τα 70 είναι της Lady Gaga και τα 30 του Βασίλη Καρρά.

Έστω πως ακούμε τρία τραγούδια στην τύχη. Παρατηρούμε πως το πλήθος X των τραγουδιών της Lady Gaga που επιλέξαμε έχει κατανομή $X \sim \text{Διων}(3, 70/100)$. Συνεπώς, η πιθανότητα να επιλέξαμε τουλάχιστον 2 τραγούδια της Lady Gaga είναι

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - p_X(0) - p_X(1) \\ &= 1 - \binom{3}{0}(0.7)^0(0.3)^3 - \binom{3}{1}(0.7)^1(0.3)^2 \simeq 0.784. \end{aligned}$$

Το μέσο πλήθος τραγουδιών της Lady Gaga που επιλέξαμε είναι $E(X) = 3 \times \frac{70}{100} = 2.1$.

Τέλος, αν συνεχίζαμε να ακούμε τραγούδια (πάλι με επανατοποθέτηση) μέχρι την πρώτη φορά που ακούγαμε της Lady Gaga, τότε το συνολικό πλήθος, έστω Z, των επιλεγμένων τραγουδιών θα είχε κατανομή $Z \sim \text{Γεωμ}(70/100)$, και κατά μέσο όρο θα επιλέξουμε $E(Z) = \frac{100}{70}$ τραγούδια. Επιπλέον, από την ιδιότητα έλλειψης μνήμης έχουμε

$$\begin{aligned} &P(\text{Ακούσαμε } 5 | \text{Ακούσαμε πάνω από } 3) \\ &\qquad = P(Z = 2 + 3 | Z > 3) = P(Z = 2) = \left(1 - \frac{70}{100}\right) \times \frac{70}{100} = 0.21. \end{aligned}$$

Παράδειγμα 5.6. *(Χαλασμένο τηλέφωνο)* Σε μια τηλεφωνική συνομιλία, κατά τη διάρκεια κάθε λεπτού υπάρχει μια πιθανότητα 0.05 να «πέσει» η γραμμή. Υποθέτουμε πως η συμπεριφορά της τηλεφωνικής γραμμής από λεπτό σε λεπτό είναι ανεξάρτητη. Θα απαντήσουμε τα ακόλουθα:

1. Ποια η πιθανότητα να πέσει η γραμμή για πρώτη φορά κατά το 5ο λεπτό της συνομιλίας;

2. Ποια η πιθανότητα να μην έχει πέσει η γραμμή κατά τα πρώτα 10 λεπτά της συνομιλίας;

3. Με δεδομένο ότι η γραμμή δεν έχει πέσει το πρώτο μισάωρο, ποια είναι η πιθανότητα να πέσει το 35ο λεπτό; Ποια η πιθανότητα να μην έχει πέσει κατά τα πρώτα 40 λεπτά;

Έστω X Τ.Μ. που δείχνει σε ποιο λεπτό της συνομιλίας έπεσε η γραμμή. Η X ακολουθεί την γεωμετρική κατανομή με παράμετρο επιτυχίας $p = 0.05$, αφού κάθε λεπτό αποτελεί ανεξάρτητη δοκιμή Bernoulli με πιθανότητα επιτυχίας $p = 0.05$. Εδώ, επιτυχία σημαίνει να «πέσει» η γραμμή.

1. Η ζητούμενη πιθανότητα είναι
$$P(X = 5) = (1 - 0.05)^4 \times 0.05 \simeq 0.0407.$$

2. Η γραμμή δεν θα πέσει τα πρώτα 10 λεπτά αν και μόνο αν έχουμε 10 αποτυχίες στη σειρά, δηλαδή
$$P(X > 10) = (1 - 0.05)^{10} \simeq 0.5987.$$

3. Χρησιμοποιώντας την ιδιότητα της απώλειας μνήμης, άμεσα προκύπτει ότι οι ζητούμενες πιθανότητες ισούνται με αυτές των πρώτων δύο σκελών. Πιο αναλυτικά:
$$\begin{aligned} P(X = 35|X > 30) &= P(X = 5) = (1 - 0.05)^4 \times 0.05 \simeq 0.0407, \\ P(X > 40|X > 30) &= P(X > 10) = (1 - 0.05)^{10} \simeq 0.5987. \end{aligned}$$

Παράδειγμα 5.7. *(Υπολογισμός παραμέτρου γεωμετρικής κατανομής)* Έστω ότι μια Τ.Μ. X έχει γεωμετρική κατανομή, και γνωρίζουμε ότι η πιθανότητα $P(X < 5) = 50\%$. Με βάση αυτό το στοιχείο, θα βρούμε τη μέση τιμή και τη διασπορά της.

Έστω p η παράμετρος της X.
$$\begin{aligned} P(X < 5) = 0.5 \Rightarrow P(X \geq 5) = 0.5 \Rightarrow (1 - p)^4 = 0.5 \\ \Rightarrow 1 - p = 0.5^{1/4} \Rightarrow p = 1 - 0.5^{1/4} \simeq 0.1591. \end{aligned}$$

Άρα
$$E(X) = \frac{1}{p} = \frac{1}{1 - 0.5^{1/4}} \simeq 6.2825, \ \ \mathrm{VAR}(X) = \frac{1-p}{p^2} = \frac{1 - (1 - 0.5^{1/4})}{(1 - 0.5^{1/4})^2} \simeq 33.2187.$$

Παράδειγμα 5.8. *(Δύο επιτυχίες)* Ένα πείραμα με πιθανότητα επιτυχίας p εκτελείται διαδοχικά, μέχρι να έρθει η δεύτερη επιτυχία. Έστω X η προσπάθεια στην οποία θα έρθει η δεύτερη επιτυχία. Θα προσδιορίσουμε την κατανομή $p_X(x)$ της X.

Παρατηρούμε καταρχήν ότι $S_X = \{2, 3, \dots\}$. Έστω A το ενδεχόμενο να έχουμε ακριβώς μια επιτυχία στις πρώτες $x - 1$ προσπάθειες, και έστω B το ενδεχόμενο να έχουμε επιτυχία στην προσπάθεια x. Παρατηρήστε πως $X = x$ ανν συμβούν και τα δύο αυτά ενδεχόμενα, τα οποία επιπλέον είναι ανεξάρτητα. Επομένως,
$$\begin{aligned} p_X(x) &= P(X = x) = P(AB) = P(A)P(B) \\ &= \binom{x-1}{1} p(1-p)^{x-1} p = \binom{x-1}{1} p^2 (1-p)^{x-2} = (x-1)p^2(1-p)^{x-2}. \end{aligned}$$

Η τέταρτη ισότητα προέκυψε με χρήση της διωνυμικής κατανομής.

5.5 Υπεργεωμετρική Κατανομή

Λήμμα 5.7. (Δειγματοληψία χωρίς επανάθεση) *Έστω πως*

$$\text{από } N \text{ αντικείμενα όπου } \begin{cases} \text{τα } k \text{ είναι τύπου I,} \\ \text{τα } N-k \text{ είναι τύπου II,} \end{cases}$$

επιλέγουμε τυχαία n, χωρίς επανατοποθέτηση, και χωρίς να υπάρχει κάποια προτίμηση στην επιλογή ορισμένων εξ αυτών. Υποθέτουμε πως

$$N \in \{1, 2, \dots\}, \quad k \in \{0, 1, \dots, N\}, \quad n \in \{1, 2, \dots, N\}.$$

Ο αριθμός X των αντικειμένων τύπου I που βρίσκονται στα n που επιλέξαμε λαμβάνει τιμές στο σύνολο των ακεραίων $S_X = \{\max(0, n-(N-k)), \dots, \min\{k, n\}\}$, και επιπλέον, αν $x \in S_X$, τότε

$$P(X = x) = \binom{k}{x}\binom{N-k}{n-x} \Big/ \binom{N}{n}.$$

Απόδειξη. Ο αριθμός x από αντικείμενα τύπου I που επιλέγουμε πρέπει να είναι το πολύ ίσος με τον αριθμό k των αντικειμένων τύπου I που υπάρχουν και τον συνολικό αριθμό n αντικειμένων που επιλέγουμε. Επιπλέον, δεν μπορεί να είναι αρνητικός. Τέλος, αν $n > N - k$, δηλαδή επιλέξαμε περισσότερα αντικείμενα από τα αντικείμενα τύπου II, είναι βέβαιο πως πήραμε τουλάχιστον $n - (N - k)$ τύπου I. Άρα, το S_X είναι ως άνω.

Έστω τώρα $x \in S_X$. Παρατηρούμε πως ο δειγματικός χώρος περιλαμβάνει συνολικά $\binom{N}{n}$ συνδυασμούς. Για να υπολογίσουμε το πλήθος των συνδυασμών που καταλήγουν σε x αντικείμενα τύπου I, παρατηρούμε πως έχουμε $\binom{k}{x}$ επιλογές για το ποια αντικείμενα τύπου I θα επιλέξουμε, και $\binom{N-k}{n-x}$ επιλογές για τα αντικείμενα τύπου II. Από την πολλαπλασιαστική αρχή, προκύπτει πως υπάρχουν τελικά $\binom{k}{x}\binom{N-k}{n-x}$ συνδυασμοί με x αντικείμενα τύπου I, και η πιθανότητα να προκύψουν x αντικείμενα είναι ως άνω. □

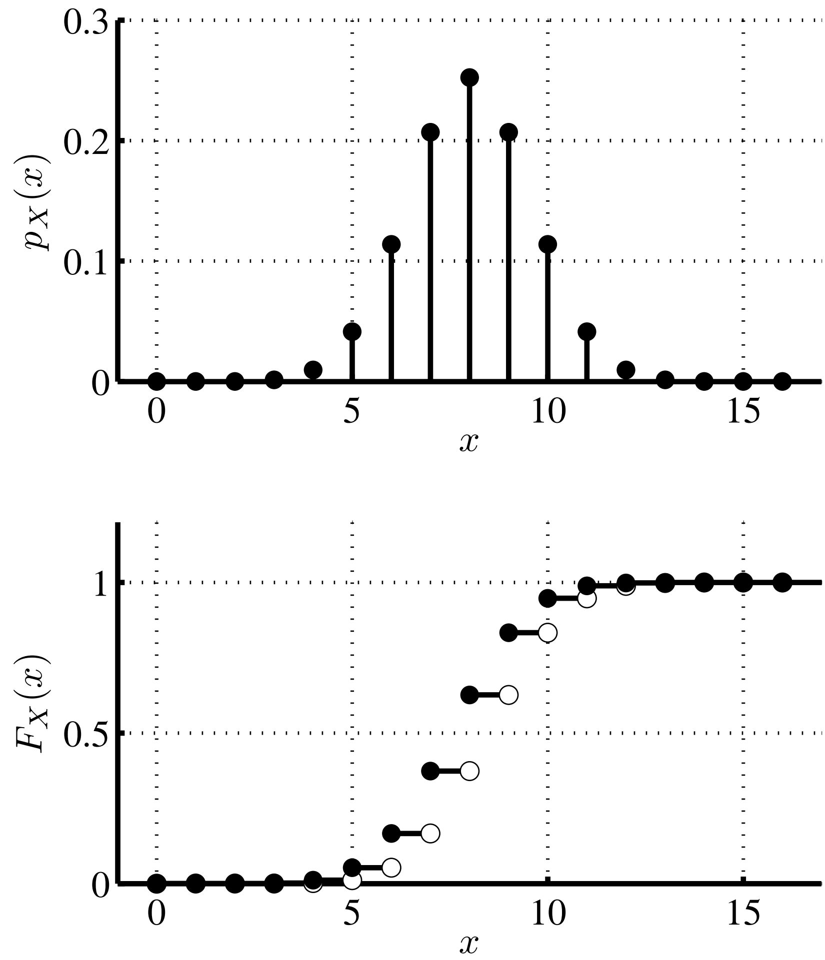
Ορισμός 5.4. (Υπεργεωμετρική κατανομή) *Μια διακριτή Τ.Μ. X ακολουθεί την <u>υπεργεωμετρική κατανομή με παραμέτρους N, k και n</u>, όπου*

$$N \in \{1, 2, \dots\}, \quad k \in \{0, 1, \dots, N\}, \quad n \in \{1, 2, \dots, N\},$$

αν έχει μάζα και σύνολο τιμών τα ακόλουθα:

$$p_X(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \quad x \in S_X = \{\max(0, n-(N-k)), \dots \min\{k, n\}\}.$$

Για συντομία, γράφουμε $X \sim \text{Υπερ}(N, k, n)$.



Σχήμα 5.5: Η μάζα και η κατανομή μιας υπεργεωμετρικής Τ.Μ. με $N = 40$, $k = 20$, $n = 16$.

Παράδειγμα 5.9. Έστω πως επιλέγουμε τυχαία τρία φύλλα από μια συνηθισμένη τράπουλα. Ποια είναι η πιθανότητα να επιλέξαμε ακριβώς μία φιγούρα; Αν ορίσουμε την Τ.Μ. X ως το πλήθος από φιγούρες που επιλέξαμε, τότε (αφού η τράπουλα περιέχει 12 φιγούρες) η X έχει κατανομή Υπερ$(52, 12, 3)$ και η ζητούμενη πιθανότητα ισούται με

$$P(X = 1) = p_X(1) = \frac{\binom{12}{1}\binom{52-12}{3-1}}{\binom{52}{3}} \simeq 0.4235.$$

Παράδειγμα 5.10. *(Mini exit poll)* Σε ένα σύλλογο που αποτελείτε από 100 άτομα, κατατίθεται προς ψήφιση μια πρόταση υπέρ της οποίας ψηφίζουν 60 άτομα. Προ της καταμέτρησης, ερωτώνται τυχαία 5 άτομα τι ψήφισαν. Ποια η πιθανότητα η πλειοψηφία αυτών να ήταν κατά της πρότασης, και έτσι να αποκτήσουμε τη λάθος εντύπωση για το τελικό αποτέλεσμα;

Για να απαντήσουμε το άνω ερώτημα, έστω X το πλήθος των ατόμων που ερωτήθηκαν

και ψήφισαν κατά. Παρατηρούμε πως $X \sim$ Υπερ$(100, 40, 5)$, και επομένως

$$\begin{aligned} P(X \geq 3) &= P(X=3) + P(X=4) + P(X=5) \\ &= \frac{\binom{40}{3}\binom{100-40}{5-3}}{\binom{100}{5}} + \frac{\binom{40}{4}\binom{100-40}{5-4}}{\binom{100}{5}} + \frac{\binom{40}{5}\binom{100-40}{5-5}}{\binom{100}{5}} \simeq 0.3139. \end{aligned}$$

Λήμμα 5.8. *Έστω $X \sim$ Υπερ(N, k, n). Η μέση τιμή της ισούται με $E(X) = nk/N$ και η διασπορά της ισούται με $\mathrm{VAR}(X) = [nk(N-k)(N-n)]/[N^2(N-1)]$.*

Παρατήρηση: Με την θεωρία που έχουμε αναπτύξει μέχρι τώρα, η εύρεση της μέσης τιμής και της διασποράς είναι χρονοβόρα, και προς το παρόν αναβάλλεται.

5.6 Κατανομή Poisson

Σε αντίθεση με τις κατανομές που έχουμε δει μέχρι τώρα, η κατανομή Poisson δεν προκύπτει άμεσα από κάποιο μοντέλο σύνθετου πειράματος. Όμως είναι εξαιρετικά χρήσιμη γιατί:

1. Εμφανίζεται ως όριο της διωνυμικής όταν έχουμε πολύ μεγάλο αριθμό πειραμάτων με πολύ μικρή πιθανότητα επιτυχίας το καθένα.

2. Επιπλέον, έχει ορισμένες πολύ χρήσιμες ιδιότητες που την καθιστούν εξαιρετικά χρήσιμη στην θεωρητική ανάλυση αλγορίθμων. Μερικές από αυτές τις ιδιότητες θα τις δούμε σύντομα.

Σημειωτέον, υπάρχουν πολλά παραδείγματα περιπτώσεων όπου έχουμε πολύ μεγάλο αριθμό πειραμάτων με μικρή πιθανότητα επιτυχίας. Για παράδειγμα:

1. Κάθε μέρα πετούν πολλά αεροπλάνα, κάθε ένα από τα οποία θα πέσει με πολύ μικρή πιθανότητα.

2. Κάθε άνθρωπος παίρνει στη διάρκεια της ζωής του πολλές φορές το αυτοκίνητο, και η πιθανότητα να τρακάρει κάθε φορά είναι πολύ μικρή.

3. Επισκεπτόμαστε το Internet πολλές φορές στη διάρκεια ενός χρόνου, αλλά η πιθανότητα να επισκεφτούμε το www.expedia.com είναι πολύ μικρή κάθε φορά.

4. Υπάρχουν πολλοί χρήστες στο Internet, αλλά ο καθένας τους θα επισκεφτεί την σελίδα μας με πολύ μικρή πιθανότητα.

Ορισμός 5.5. (Κατανομή Poisson) *Μια διακριτή τυχαία μεταβλητή X ακολουθεί την κατανομή Poisson με παράμετρο $\lambda > 0$, $\lambda \in \mathbb{R}$, όταν το σύνολο τιμών της είναι το $S_X = \{0, 1, 2, \dots\}$ και η μάζα της ισούται με*

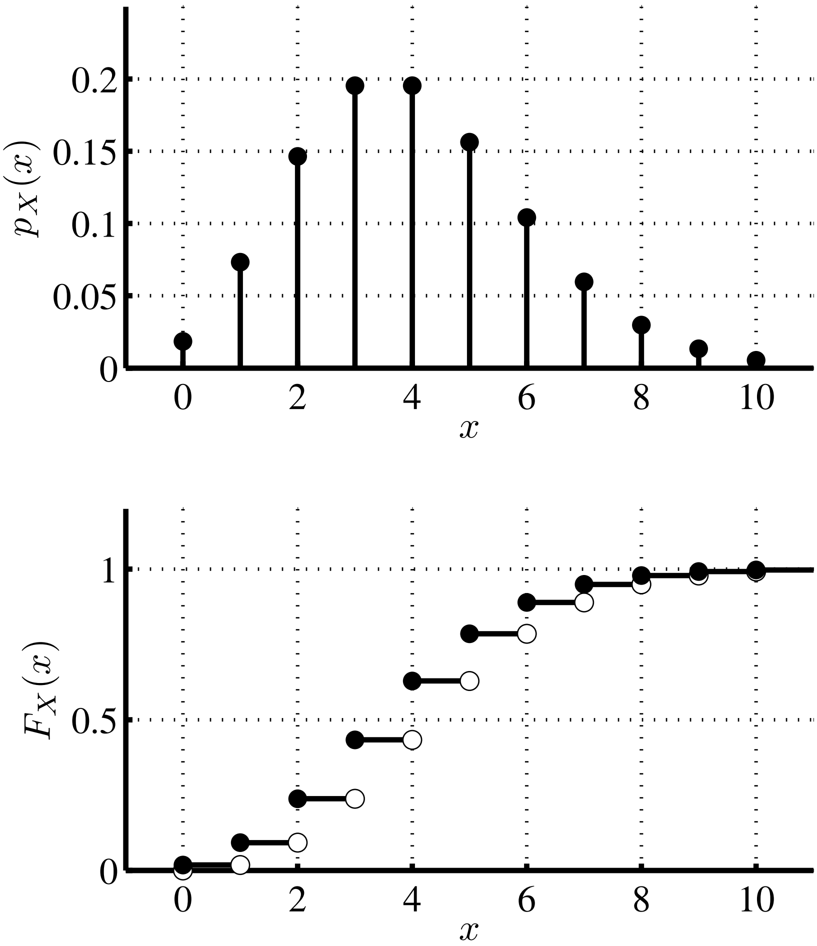
$$p_X(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x \in S_X = \{0, 1, 2, \dots\}.$$

Γράφουμε $X \sim \text{Poisson}(\lambda)$.

Παρατηρήσεις

1. Παρατηρήστε πως

$$\sum_{x=0}^{\infty} p_X(x) = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!} = e^{-\lambda} \times e^{\lambda} = 1,$$



Σχήμα 5.6: Η μάζα και η κατανομή μιας Τ.Μ. Poisson με παράμετρο $\lambda = 4$.

όπως πρέπει, για να είναι η $p_X(x)$ μάζα. Χρησιμοποιήσαμε τη γνωστή δυναμοσειρά

$$e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}, \qquad \lambda \in \mathbb{R}. \tag{5.5}$$

2. Στο Σχήμα 5.6 έχουμε σχεδιάσει την μάζα και την κατανομή μιας Τ.Μ. Poisson με παράμετρο $\lambda = 4$.

Θεώρημα 5.1. (Η κατανομή Poisson προσεγγίζει την διωνυμική κατανομή) *Έστω διωνυμική Τ.Μ. Y με παραμέτρους N και $p = \frac{\lambda}{N}$, όπου $\lambda > 0$. Καθώς το N τείνει στο άπειρο, η μάζα της Y συγκλίνει στην μάζα μιας Τ.Μ. X με κατανομή Poisson και παράμετρο λ, δηλαδή*

$$p_Y(m) = \binom{N}{m} \left(\frac{\lambda}{N}\right)^m \left(1 - \frac{\lambda}{N}\right)^{N-m} \to p_X(m) = e^{-\lambda}\frac{\lambda^m}{m!},$$

καθώς το $N \to \infty$.

Απόδειξη. Παρατηρούμε πως:

$$\begin{aligned} p_Y(m) &= \binom{N}{m}\Big(\frac{\lambda}{N}\Big)^m\Big(1-\frac{\lambda}{N}\Big)^{N-m} = \frac{N!\lambda^m}{(N-m)!m!N^m}\left(1-\frac{\lambda}{N}\right)^{N-m} \\ &= \frac{N(N-1)\dots(N-m+1)}{N^m}\times\left(1-\frac{\lambda}{N}\right)^{N}\times\left(1-\frac{\lambda}{N}\right)^{-m}\times\frac{\lambda^m}{m!}. \end{aligned}$$

Το άνω γινόμενο περιλαμβάνει 4 όρους, εκ των οποίων ο πρώτος τείνει στο 1, καθώς είναι ρητή συνάρτηση με ίσους μεγιστοβάθμιους όρους στον αριθμητή και τον παρονομαστή. Ο δεύτερος τείνει στο $e^{-\lambda}$, σύμφωνα με γνωστό όριο, ενώ ο τρίτο τείνει προφανώς στη μονάδα. Άρα τελικά προκύπτει το ζητούμενο. □

Παρατηρήσεις

1. Η ύπαρξη αυτής της προσέγγισης είναι χρήσιμη για δύο λόγους. Πρώτον, γιατί σε ορισμένες περιπτώσεις μειώνει σημαντικά τον αριθμό των πράξεων. Δεύτερον, γιατί μας επιτρέπει, σε ορισμένες περιπτώσεις που εμφανίζεται η διωνυμική κατανομή, να εισάγουμε την κατανομή Poisson και να χρησιμοποιήσουμε τις πολύ καλές τις ιδιότητες.

2. Πρακτικά, όταν έχουμε να υπολογίσουμε τη μάζα μιας διωνυμικής Τ.Μ. όπου ο αριθμός των πειραμάτων είναι πολύ μεγάλος, π.χ., μεγαλύτερος του 100, η πιθανότητα επιτυχίας του κάθε πειράματος p είναι πολύ μικρή, π.χ. μικρότερη του 1/25, και το γινόμενο Np είναι της τάξης του 1, τότε μπορούμε να χρησιμοποιούμε την μάζα της κατανομής Poisson αντί για τη μάζα της διωνυμικής κατανομής. Δείτε το επόμενο παράδειγμα.

Παράδειγμα 5.11. *(Επιθέσεις μηνυμάτων spam)* Ένας Νιγηριανός που μόλις κληρονόμησε 10^8 ευρώ στέλνει ένα email στους 10 χιλιάδες λογαριασμούς του domain aueb.gr, ζητώντας τον αριθμό της πιστωτικής κάρτας του παραλήπτη. Από προηγούμενες προσπάθειές του γνωρίζει ότι οι διαφορετικοί χρήστες ανταποκρίνονται ανεξάρτητα ο ένας απ᾽ τον άλλον, και η πιθανότητα να απαντήσει κάποιος με τα ζητούμενα στοιχεία είναι 0.018%. Ποια η πιθανότητα να του απαντήσουν τουλάχιστον τρεις χρήστες;

Έστω X το πλήθος των χρηστών που ανταποκρίνονται στο email. Το X εκφράζει το πλήθος των «επιτυχιών» σε $N = 10$ χιλιάδες όμοια, ανεξάρτητα πειράματα, που το καθένα έχει πιθανότητα επιτυχίας $p = 0.00018$. Συνεπώς $X \sim \text{Διων}(10000, 0.00018)$.

Επειδή το γινόμενο $Np = 10000 \times 0.00018 = 1.8$ είναι της τάξεως του 1, μπορούμε να προσεγγίσουμε την κατανομή της X μέσω της Poisson(λ), με $\lambda = Np = 10000 \times 0.00018 = 1.8$. Από τον τύπο της μάζας της κατανομής Poisson έχουμε

$$\begin{aligned} P(X\geq 3) &= 1-P(X\leq 2) = 1-P(X=0)-P(X=1)-P(X=2) \\ &\simeq 1-e^{-1.8}\times\frac{1.8^0}{0!}-e^{-1.8}\times\frac{1.8^1}{1!}-e^{-1.8}\times\frac{1.8^2}{2!}\simeq 0.2693789. \end{aligned}$$

Δεδομένου του μηδαμινού κόστους της αποστολής των μηνυμάτων, η πρόβλεψη για τον κληρονόμο δεν είναι κι άσχημη!

Χωρίς την προσέγγιση, θα είχαμε:

$$\begin{aligned} P(X \geq 3) &= 1 - P(X \leq 2) = 1 - P(X=0) - P(X=1) - P(X=2) \\ &= 1 - \binom{10000}{0}(1-0.00018)^{10000} - \binom{10000}{1}(1-0.00018)^{9999}0.00018 \\ &\quad - \binom{10000}{2}(1-0.00018)^{9998}0.00018^2 \simeq 0.2693741. \end{aligned}$$

Άρα σε αυτή την περίπτωση η προσέγγιση είναι εξαιρετικά ακριβής.

Παράδειγμα 5.12. *(Συλλέκτης κουπονιών)* Έστω πως έχουμε αγοράσει από το περίπτερο, κατά τη διάρκεια του χρόνου, 200 φακέλους που περιέχουν εικόνες ποδοσφαιριστών. Υπάρχουν συνολικά 100 εικόνες, και κάθε φάκελος είναι εξίσου πιθανό να περιέχει μια οποιαδήποτε από αυτές. Ποια η πιθανότητα να έχω i φορές την εικόνα ενός συγκεκριμένου ποδοσφαιριστή;

Για να απαντήσουμε το ερώτημα, φανταστείτε πως έχουμε 200 ανεξάρτητα πειράματα, καθένα με πιθανότητα επιτυχίας 1/100. Άρα, ο αριθμός των φορών X που θα έχουμε επιτυχία δίνεται από την διωνυμική κατανομή με παραμέτρους $N = 200$, $p = 0.01$. Λόγω της επιλογής όμως των παραμέτρων, μπορώ να χρησιμοποιήσω την προσέγγιση Poisson. Δείτε τον παρακάτω πίνακα:

x	$p_X(x)$, Διωνυμική	$p_X(x)$, Poisson
0	0.1340	0.1353
1	0.2707	0.2707
2	0.2720	0.2707
3	0.1814	0.1804
4	0.0902	0.0902

Λήμμα 5.9. (Ιδιότητες κατανομής Poisson) *Έστω Τ.Μ.* $X \sim \text{Poisson}(\lambda)$.

1. $E(X) = \lambda$.
2. $\text{VAR}(X) = \lambda$.

Απόδειξη. 1.

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x e^{-\lambda}\frac{\lambda^x}{x!} = \sum_{x=1}^{\infty} x e^{-\lambda}\frac{\lambda^x}{x!} \\ &= e^{-\lambda}\lambda\sum_{x=1}^{\infty}\frac{\lambda^{x-1}}{(x-1)!} = e^{-\lambda}\lambda\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = e^{-\lambda}\lambda e^{\lambda} = \lambda. \end{aligned}$$

Στην τέταρτη ισότητα κάναμε την αντικατάσταση $j = x - 1$. Στην πέμπτη, κάναμε χρήση της (5.5).

2. Με παρόμοιο τρόπο, καταρχήν υπολογίζουμε το $E(X(X-1))$:

$$\begin{aligned} E(X(X-1)) &= \sum_{x=0}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} = \sum_{x=2}^{\infty} x(x-1)e^{-\lambda}\frac{\lambda^x}{x!} \\ &= e^{-\lambda}\lambda^2 \sum_{x=2}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} = e^{-\lambda}\lambda^2 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = e^{-\lambda}\lambda^2 e^{\lambda} = \lambda^2. \end{aligned}$$

Στην τέταρτη ισότητα κάναμε την αντικατάσταση $j = x - 2$. Στην πέμπτη, κάναμε χρήση της (5.5). Ακολούθως, έχουμε:

$$\begin{aligned} E(X^2) &= E(X(X-1)) + E(X) = \lambda^2 + \lambda, \\ \mathrm{VAR}(X) &= E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \end{aligned}$$

□

5.7 Παραδείγματα

Παράδειγμα 5.13. *(Διάφορες κατανομές)* Παρακάτω ορίζονται κάποιες Τ.Μ. Για την κάθε μία, θα περιγράψουμε την κατανομή της και τις αντίστοιχες παραμέτρους.

1. Ρίχνουμε διαδοχικές (ανεξάρτητες) ζαριές με δύο ζάρια, μέχρι την 1^{η} φορά που θα φέρουμε διπλή. Έστω X το συνολικό πλήθος από ζαριές που ρίξαμε.

2. Επιλέγουμε στην τύχη, με επανατοποθέτηση, 6 φύλλα από μια τράπουλα. Έστω X το πλήθος από κούπες που επιλέξαμε.

3. Όπως στο (β΄), αλλά χωρίς επανατοποθέτηση.

4. Ρίχνουμε 20 φορές ένα κέρμα με πιθανότητα κορώνας $P(H) = 0.3$. Έστω X το πλήθος των φορών που φέραμε γράμματα.

Έχουμε, κατά περίπτωση:

1. Η X ακολουθεί τη γεωμετρική κατανομή με παράμετρο $p = 1/6$ αφού η πιθανότητα να φέρουμε διπλή είναι $1/6$, και τα πειράματα (δηλαδή οι διαδοχικές ζαριές) είναι ανεξάρτητα μεταξύ τους.

2. Η X ακολουθεί τη διωνυμική κατανομή με παραμέτρους $N = 6$ και $p = 1/4$, αφού έχουμε το πλήθος επιτυχιών σε $N = 6$ όμοια, ανεξάρτητα πειράματα με πιθανότητα επιτυχίας $p = 13/52 = 1/4$.

3. Η X ακολουθεί την υπεργεωμετρική κατανομή με παραμέτρους $(52, 13, 6)$ αφού επιλέγουμε χωρίς επανατοποθέτηση 6 φύλλα απ΄ τα 52 που έχει συνολικά μία τράπουλα, και εκ των οποίων τα 13 είναι κούπες.

4. Η X ακολουθεί την διωνυμική κατανομή με παραμέτρους $N = 20$ και $p = 0.7$, αφού X είναι το πλήθος επιτυχιών σε $N = 20$ όμοια, ανεξάρτητα πειράματα με πιθανότητα επιτυχίας την πιθανότητα να φέρουμε γράμματα $P(T) = 1 - 0.3 = 0.7$.

Παράδειγμα 5.14. *(Ουρά σε τράπεζα)* Σε μια τράπεζα, κάθε πεντάλεπτο υπάρχει πιθανότητα 5% να έρθει ένας νέος πελάτης. Υποθέτουμε πως οι αφίξεις πελατών σε διαφορετικά πεντάλεπτα είναι ανεξάρτητες, και ορίζουμε X το πλήθος πελατών που έφτασαν τις πρώτες 2 ώρες και Y το πρώτο 5λεπτο κατά το οποίο έφτασε πελάτης. Αγνοούμε το ενδεχόμενο να έρθουν 2 ή περισσότεροι πελάτες σε κάποιο πεντάλεπτο. Θα απαντήσουμε στα ακόλουθα:

1. Ποια είναι η κατανομή του X;

2. Ποια πιθανότητα να έρθουν ακριβώς 3 πελάτες τις πρώτες δύο ώρες.

3. Ποια είναι η κατανομή της Y;
4. Κατά μέσο όρο πόσα πεντάλεπτα θα περιμένουν οι υπάλληλοι μέχρι την άφιξη του πρώτου πελάτη;
5. Δεδομένου ότι δεν ήρθε κανείς τις πρώτες 2 ώρες, ποια είναι η πιθανότητα να μην έρθει κανείς και κατά την επόμενη μισή ώρα;

Έχουμε, κατά περίπτωση:

1. Οι 2 ώρες έχουν 24 πεντάλεπτα, κι εφόσον η πιθανότητα επιτυχίας σε κάθε πεντάλεπτο, δηλαδή να έρθει ένας νέος πελάτης στο πεντάλεπτο, είναι 0.05, θα έχουμε $X \sim$Διων$(24, 0.05)$.
2. Η πιθανότητα να έρθουν ακριβώς 3 πελάτες τις πρώτες δύο ώρες είναι
$$P(X=3) = p_X(3) = \binom{24}{3}(0.05)^3(1-0.05)^{24-3} \simeq 0.0862.$$
3. Η Y περιγράφει τη χρονική στιγμή της πρώτης επιτυχίας, άρα έχει γεωμετρική κατανομή με παράμετρο $p = 0.05$.
4. Ο μέσος χρόνος που θα περιμένουν οι υπάλληλοι μέχρι την άφιξη του πρώτου πελάτη είναι η μέση τιμή της Y. Έχουμε: $E(Y) = 1/0.05 = 20$ πεντάλεπτα.
5. Παρατηρήστε ότι οι 2 ώρες έχουν 24 πεντάλεπτα και οι 2.5 ώρες έχουν 30 πεντάλεπτα. Λόγω της ιδιότητας έλλειψης μνήμης έχουμε:
$$\begin{aligned} P(Y > 30 | Y > 24) &= P(Y > 24 + 6 | Y > 24) = P(Y > 6) \\ &= (1-0.05)^6 \simeq 0.7351. \end{aligned}$$

Παράδειγμα 5.15. *(Τέσσερα ζάρια)* Παίζουμε το εξής παιχνίδι: Ρίχνουμε τέσσερα ζάρια συγχρόνως, και κερδίζουμε όταν έρθουν 6 και τα τέσσερα. Επαναλαμβάνουμε το ίδιο πείραμα 1000 φορές. Θα υπολογίσουμε την πιθανότητα στα 1000 παιχνίδια να κερδίσαμε ακριβώς δύο φορές.

Έστω X το πλήθος των φορών που κερδίσαμε. Η κατανομή της X είναι διωνυμική, με παραμέτρους $N = 1000$ και $p = \left(\frac{1}{6}\right)^4 = 1/1296$. Άρα,
$$P(X=2) = \binom{1000}{2}\left(\frac{1}{1296}\right)^2\left(1-\frac{1}{1296}\right)^{998} \simeq 0.137645.$$
Εναλλακτικά, παρατηρούμε ότι το N είναι αρκετά μεγάλο, το p είναι αρκετά μικρό, και $Np = \frac{1000}{1296} \simeq 0.7716$. Συνεπώς, μπορούμε να προσεγγίσουμε την κατανομή με την κατανομή Poisson με παράμετρο $\lambda = 0.7716$. Με βάση αυτή την προσέγγιση, έχουμε:
$$P(X=2) \simeq \exp(-0.7716) \times \frac{0.7716^2}{2!} \simeq 0.137611.$$

Όνομα	Σύνολο τιμών και μάζα	Μέση τιμή	Διασπορά	Φυσική ερμηνεία της X
Bernoulli	$S_X = \{0, 1\}$ $p(0) = 1 - p, \quad p(1) = p$	$E(X) = p$	$\mathrm{VAR}(X) = p(1-p)$	Αριθμός επιτυχιών σε 1 πείραμα με πιθανότητα επιτυχίας p.
Διωνυμική	$S_X = \{0, 1, \dots, N\}$ $p_X(x) = \binom{N}{x} p^x (1-p)^{N-x}$	$E(X) = Np$	$\mathrm{VAR}(X) = Np(1-p)$	Αριθμός επιτυχιών σε N όμοια και ανεξάρτητα πειράματα με πιθανότητα επιτυχίας p.
Γεωμετρική	$S_X = \{1, 2, \dots\}$ $p_X(x) = (1-p)^{x-1} p$	$E(X) = 1/p$	$\mathrm{VAR}(X) = (1-p)/p^2$	Αριθμός προσπαθειών μέχρι την πρώτη επιτυχία σε ακολουθία ανεξάρτητων όμοιων πειραμάτων με πιθανότητα επιτυχίας p.
Υπεργεωμετρική	$S_X = \{\max\{0, n-(N-k)\} \dots, \min\{k, n\}\}$ $p_X(x) = \binom{k}{x}\binom{N-k}{n-x}/\binom{N}{n}$	$E(X) = nk/N$	$\mathrm{VAR}(X) = \frac{nk(N-k)(N-n)}{N^2(N-1)}$	Επιλέγουμε n αντικείμενα από N εκ των οποίων k είναι τύπου Ι. Το X είναι το πλήθος των αντικειμένων τύπου Ι από τα n που επιλέξαμε.
Poisson	$S_X = \{0, 1, \dots\}$ $p_X(x) = e^{-\lambda}\lambda^x / x!$	$E(X) = \lambda$	$\mathrm{VAR}(X) = \lambda$	Προσέγγιση της διωνυμικής όταν $p = \lambda/N$ και $N \to \infty$.

Πίνακας 5.1: Συνήθεις διακριτές Τ.Μ.

Κεφάλαιο 6

Ζεύγη Διακριτών Τυχαίων Μεταβλητών

Η έννοια της συνάρτησης μάζας πιθανότητας μας επιτρέπει την μεμονωμένη μελέτη Τ.Μ. Σε προβλήματα όμως όπου εμφανίζονται πολλές Τ.Μ, η μεμονωμένη μελέτη κάθε μιας από αυτές δεν αποκαλύπτει τα πάντα. Φανταστείτε για παράδειγμα πως κάποιος ρίχνει δύο ζάρια με τέτοιο τρόπο ώστε

$$P((1,1)) = P((2,2)) = P((3,3)) = P((4,4)) = P((5,5)) = P((6,6)) = \frac{1}{6},$$

ενώ όλα τα άλλα ενδεχόμενα έχουν μηδενική πιθανότητα. Εύκολα μπορούμε να υπολογίσουμε ότι, μεμονωμένα, κάθε ένα από τα ζάρια είναι «δίκαιο», αφού η πιθανότητα να προκύψει οποιοδήποτε νούμερο είναι $\frac{1}{6}$. Ο συνδυασμός των ρίψεων, όμως, ΔΕΝ είναι «δίκαιος», και αυτό είναι κάτι που δεν μας αποκαλύπτουν οι μάζες πιθανοτήτων καθενός από τα δύο ζάρια! Το κενό αυτό καλύπτεται από την από κοινού συνάρτηση μάζας πιθανότητας, την συνδιακύμανση, και άλλες έννοιες που θα δούμε σε αυτό το κεφάλαιο.

6.1 Από Κοινού Μάζα

Ορισμός 6.1. (Από κοινού μάζα και περιθώρια μάζα) *Έστω δύο διακριτές Τ.Μ. X, Y, με σύνολα τιμών S_X και S_Y αντιστοίχως.*

1. *Ορίζουμε την <u>από κοινού συνάρτηση μάζας πιθανότητας</u> ή <u>από κοινού μάζα</u> ως την συνάρτηση $p_{XY}(x,y) : S_X \times S_Y \to [0,1]$ που ορίζεται ως:*

$$p_{XY}(x,y) = P(X = x, Y = y), \quad \forall x \in S_X, \ \forall y \in S_Y.$$

2. *Οι μάζες πιθανότητας των X, Y καλούνται <u>περιθώριες μάζες πιθανότητας</u>.*

Παράδειγμα 6.1. *(Δύο προγράμματα)* Δύο διαφορετικά προγράμματα κατανέμονται τυχαία σε τρεις υπολογιστές, χωρίς κάποιο από αυτά να δείχνει προτίμηση σε κάποιον

υπολογιστή, και ανεξάρτητα το ένα από το άλλο. Έστω X, Y, Z τα πλήθη των προγραμμάτων που κατέληξαν στον κάθε υπολογιστή, άρα $X + Y + Z = 2$. Ας εξετάσουμε τις δύο μεταβλητές X, Y. Η πιθανότητα να έχουμε $X = 0$ και $Y = 0$ είναι η πιθανότητα του ενδεχόμενου και τα δύο προγράμματα να πήγαν στον τρίτο υπολογιστή:

$$p_{XY}(0,0) = P(X = 0, Y = 0) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}.$$

Παρομοίως, η πιθανότητα να έχουμε $X = 0$ και $Y = 1$ είναι η πιθανότητα το πρώτο πρόγραμμα να πήγε στον υπολογιστή 2 και το δεύτερο στον 3, ή αντίστροφα, δηλαδή,

$$p_{XY}(0,1) = P(X = 0, Y = 1) = \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{3} = \frac{2}{9}.$$

Με τον ίδιο τρόπο μπορούν να υπολογιστούν όλες οι τιμές της p_{XY}, που συνοψίζονται στον ακόλουθο πίνακα:

x / y	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

Έστω, τώρα, πως θέλουμε να υπολογίσουμε από τις άνω τιμές τις τιμές της περιθώριας μάζας πιθανότητας της X. Χρησιμοποιώντας τον κανόνα της ολικής πιθανότητας έχουμε

$$\begin{aligned} p_X(1) &= P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) + P(X=1, Y=2) \\ &= p_{XY}(1,0) + p_{XY}(1,1) + p_{XY}(1,2) = \frac{2}{9} + \frac{2}{9} + 0 = \frac{4}{9}. \end{aligned}$$

Άρα, για να βρούμε την $p_X(1)$, αθροίσαμε όλες τις τιμές του πίνακα που αντιστοιχούσαν στην στήλη $X = 1$. Με παρόμοιο τρόπο, αθροίζοντας τις τιμές στις στήλες $X = 0$ και $X = 2$ βρίσκουμε τις $p_X(0)$ και $p_X(2)$. Αντίστοιχα, αθροίζοντας τις τιμές που βρίσκονται στην κάθε γραμμή $Y = y$ του πίνακα, βρίσκουμε τις πιθανότητες $p_Y(y)$. Συμπληρώνεται έτσι ο ακόλουθος πίνακας:

x / y	0	1	2	$p_Y(y)$
0	1/9	2/9	1/9	4/9
1	2/9	2/9	0	4/9
2	1/9	0	0	1/9
$p_X(x)$	4/9	4/9	1/9	

Λόγω του ότι οι τιμές των μαζών $p_X(x)$ και $p_Y(y)$ που αφορούν μόνο μία απ' τις δύο Τ.Μ. μπορούν να γραφτούν, όπως πιο πάνω, στο «περιθώριο» του πίνακα της από κοινού μάζας, οι επί μέρους μάζες τους ορίστηκαν άνω ως *περιθώριες*.

Παράδειγμα 6.2. *(Δύο ζάρια — συνέχεια)* Ρίχνουμε ένα ζάρι δύο φορές. Έστω πως οι ρίψεις είναι ανεξάρτητες και το ζάρι δίκαιο. Έστω X, Y, τα αποτελέσματα των ρίψεων. Οι τιμές της από κοινού μάζας $p_{XY}(x,y)$ των X, Y εύκολα υπολογίζονται ως εξής: για κάθε ζεύγος (x,y) όπου $x, y = 1,2,3,4,5,6$ έχουμε

$$p_{XY}(x,y) = P(X = x, Y = y) = P(X = x)P(Y = y) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}.$$

Η δεύτερη ισότητα προέκυψε λόγω ανεξαρτησίας, και η τρίτη λόγω του ότι το ζάρι είναι δίκαιο. Προκύπτει τελικά ο ακόλουθος πίνακας:

x / y	1	2	3	4	5	6	$p_Y(y)$
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
$p_X(x)$	1/6	1/6	1/6	1/6	1/6	1/6	

Παρατηρήστε ότι οι περιθώριες που προκύπτουν είναι οι αναμενόμενες.

Έστω τώρα ότι έρχονται πάντα διπλές, και όλες οι διπλές έχουν την ίδια πιθανότητα να εμφανιστούν. Η νέα από κοινού μάζα είναι η ακόλουθη:

x / y	1	2	3	4	5	6	$p_Y(y)$
1	1/6	0	0	0	0	0	1/6
2	0	1/6	0	0	0	0	1/6
3	0	0	1/6	0	0	0	1/6
4	0	0	0	1/6	0	0	1/6
5	0	0	0	0	1/6	0	1/6
6	0	0	0	0	0	1/6	1/6
$p_X(x)$	1/6	1/6	1/6	1/6	1/6	1/6	

Παρατηρήστε ότι οι περιθώριες μάζες των δύο περιπτώσεων ταυτίζονται, παρότι τα πειράματα από τα οποία προέρχονται είναι πολύ διαφορετικά.

Παράδειγμα 6.3. *(Δύο ζάρια — συνέχεια)* Ρίχνουμε ένα ζάρι δύο φορές, και υποθέτουμε ότι οι ρίψεις είναι ανεξάρτητες και ότι το ζάρι είναι δίκαιο. Έστω X, Y τα αποτελέσματα των δύο ρίψεων.

Έστω $Z = \min(X, Y)$ και $W = \max(X, Y)$. Θα υπολογίσουμε την από κοινού μάζα των Z, W, καθώς και τις περιθώριες μάζες τους. Έχουμε ένα πείραμα με $6 \times 6 = 36$ διαφορετικά αποτελέσματα. Επιπλέον, η από κοινού μάζα $p_{ZW}(z, w)$ των Z, W πρέπει να υπολογιστεί για $6 \times 6 = 36$ ζεύγη τιμών (z, w). Αρχικά παρατηρούμε πως κάποια απ' αυτά είναι αδύνατον να εμφανιστούν, π.χ., δεν μπορεί το ελάχιστο απ' τις δύο ζαριές να είναι 5 και το μέγιστο 2. Συνεπώς, $p_{ZW}(5, 2) = 0$, και γενικά $p_{ZW}(z, w) = 0$ αν $z > w$. Για να υπολογίσουμε κάθε μια από τις υπόλοιπες τιμές πρέπει να βρούμε τα αποτελέσματα που της αντιστοιχούν.

$$\begin{aligned} p_{ZW}(1,1) &= P(X = 1, Y = 1) = \frac{1}{36}, \\ p_{ZW}(1,2) &= P(X = 1, Y = 2) + P(X = 2, Y = 1) = \frac{2}{36}. \end{aligned}$$

Συνεχίζοντας με τον ίδιο τρόπο, συμπληρώνουμε τον ακόλουθο πίνακα:

w / z	1	2	3	4	5	6	$p_Z(z)$
1	1/36	2/36	2/36	2/36	2/36	2/36	11/36
2	0	1/36	2/36	2/36	2/36	2/36	9/36
3	0	0	1/36	2/36	2/36	2/36	7/36
4	0	0	0	1/36	2/36	2/36	5/36
5	0	0	0	0	1/36	2/36	3/36
6	0	0	0	0	0	1/36	1/36
$p_W(w)$	1/36	3/36	5/36	7/36	9/36	11/36	

Οι μάζες $p_W(w)$ και $p_Z(z)$ υπολογίζονται στο περιθώριο του πιο πάνω πίνακα, αθροίζοντας τις τιμές της από κοινού μάζας στην αντίστοιχη στήλη ή γραμμή.

Έχοντας την από κοινού μάζα στη διάθεσή μας, μπορούμε να υπολογίσουμε την πιθανότητα οποιουδήποτε ενδεχόμενου αφορά τις Z, W, ακόμα και δεσμευμένες πιθανότητες, εφόσον η δέσμευση είναι σε ενδεχόμενα που αφορούν τις Z, W. Για παράδειγμα, με χρήση των ορισμών της δεσμευμένης πιθανότητας και της από κοινού μάζας:

$$P(Z = 1 | W = 5) = \frac{P(Z = 1, W = 5)}{P(W = 5)} = \frac{p_{ZW}(1,5)}{p_W(5)} = \frac{2/36}{9/36}.$$

Λήμμα 6.1. (Ιδιότητες από κοινού μάζας) *Έστω δύο διακριτές Τ.Μ. X, Y με σύνολα τιμών S_X, S_Y και μάζες $p_X(x)$ και $p_Y(y)$, αντίστοιχα. Για την $p_{XY}(x, y)$ ισχύει:*

1. $\displaystyle\sum_{x\in S_X, y\in S_Y} p_{XY}(x, y) = 1.$

2. *Έστω A ένα οποιοδήποτε υποσύνολο του καρτεσιανού γινομένου $S_X \times S_Y$. Τότε*
$$P((X, Y) \in A) = \sum_{(x,y)\in A} p_{XY}(x, y).$$

3. $\displaystyle p_X(x) = \sum_{y\in S_Y} p_{XY}(x, y)$ *και* $\displaystyle p_Y(y) = \sum_{x\in S_X} p_{XY}(x, y).$

Απόδειξη. 1. Για διαφορετικά ζεύγη τιμών (x, y) τα ενδεχόμενα $\{X = x, Y = y\}$ είναι ξένα μεταξύ τους, και προφανώς η ένωση όλων τους καλύπτει όλο το Δ.Χ. Ω όπου έχουν ορισθεί, άρα από τον κανόνα ολικής πιθανότητας έχουμε
$$\begin{aligned} 1 &= P(\Omega) = P\left(\bigcup_{x\in S_X, y\in S_Y} \{X = x, Y = y\}\right) \\ &= \sum_{x\in S_X,\ y\in S_Y} P(X = x, Y = y) = \sum_{x\in S_X,\ y\in S_Y} p_{XY}(x, y). \end{aligned}$$

2. Με παρόμοιο τρόπο έχουμε
$$\begin{aligned} P((X, Y) \in A) &= P\left(\bigcup_{(x,y)\in A} \{X = x, Y = y\}\right) \\ &= \sum_{(x,y)\in A} P(X = x, Y = y) = \sum_{(x,y)\in A} p_{XY}(x, y). \end{aligned}$$

3. Για δεδομένο x και για όλα τα διαφορετικά y, τα ενδεχόμενα $\{X = x, Y = y\}$ είναι ξένα μεταξύ τους και η ένωσή τους ισούται με το ενδεχόμενο $\{X = x\}$. Άρα:
$$p_X(x) = P\left(\bigcup_{y\in S_Y} \{X = x, Y = y\}\right) = \sum_{y\in S_Y} P(X = x, Y = y) = \sum_{y\in S_Y} p_{XY}(x, y).$$
Η αντίστοιχη σχέση για την $p_Y(y)$ αποδεικνύεται με ακριβώς τον ίδιο τρόπο. (Παρατηρήστε ότι αυτό το σκέλος θα μπορούσε να προκύψει και ως ειδική περίπτωση του προηγούμενου.)

□

6.2 Παραδείγματα

Παράδειγμα 6.4. *(Συμπλήρωση πίνακα)* Οι διακριτές Τ.Μ. X και Y έχουν από κοινού μάζα που δίνεται από τον παρακάτω πίνακα:

x / y	-5	-2	2	5	$p_Y(y)$
0	0.01	0.01	0.01		
1	0.05		0.05	0	0.19
2	0.07	0.07		0.07	
3	0.15		0	0.15	
$p_X(x)$		0.26	0.15		

Ζητούνται τα ακόλουθα:

1. Να συμπληρωθούν οι τιμές που λείπουν.
2. Να υπολογιστεί η $P(Y = 1|X = 2)$.

Για το κάθε ερώτημα έχουμε:

1. Πρέπει η τιμή της περιθώριας να ισούται με το άθροισμα της αντίστοιχης γραμμής ή στήλης. Με βάση αυτή την παρατήρηση, μπορεί να συμπληρωθεί ο ακόλουθος πίνακας:

x / y	-5	-2	2	5	$p_Y(y)$
0	0.01	0.01	0.01	**0.09**	**0.12**
1	0.05	**0.09**	0.05	0	0.19
2	0.07	0.07	**0.09**	0.07	**0.30**
3	0.15	**0.09**	0	0.15	**0.39**
$p_X(x)$	**0.28**	0.26	0.15	**0.31**	

2. Από τον ορισμό της δεσμευμένης πιθανότητας και τις τιμές του πιο πάνω πίνακα έχουμε

$$P(Y = 1|X = 2) = \frac{P(Y = 1, X = 2)}{P(X = 2)} = \frac{p_{XY}(2,1)}{p_X(2)} = \frac{0.05}{0.15} = \frac{1}{3}.$$

Παράδειγμα 6.5. *(Ασανσέρ)* Σε ένα κτίριο με 3 ορόφους και ισόγειο, το ασανσέρ βρίσκεται στον όροφο X και ένας τυχαίος χρήστης στον όροφο Y όπου τα X, Y δίνονται από την ακόλουθη από κοινού μάζα πιθανότητας:

x / y	0	1	2	3	$p_Y(y)$
0	1/6	1/12	1/12	1/12	15/36
1	1/12	1/18	1/36	1/36	7/36
2	1/12	1/36	1/18	1/36	7/36
3	1/12	1/36	1/36	1/18	7/36
$p_X(x)$	15/36	7/36	7/36	7/36	

Εννοείται πως ο όροφος 0 είναι το ισόγειο.

Παρατηρήστε καταρχήν πως οι πιθανότητες του άνω πίνακα ανήκουν όλες στο $[0,1]$ και αθροίζονται στη μονάδα. Παρατηρήστε επίσης πως το άνω μοντέλο λαμβάνει υπ' όψιν ότι οι μετακινήσεις από ή προς το ισόγειο είναι πολύ συχνότερες από ότι σε άλλους ορόφους, και το ότι συχνά το ίδιο άτομο χρησιμοποιεί το ασανσέρ δύο φορές στη σειρά, στην οποία περίπτωση το ασανσέρ παραμένει στον όροφο που είναι το άτομο. Έτσι, για παράδειγμα, το ενδεχόμενο $X = 2,\ Y = 3$ είναι λιγότερο πιθανό από το $X = 2,\ Y = 2$, το οποίο με τη σειρά του είναι λιγότερο πιθανό από το $X = 2, Y = 0$.

Ακολούθως, έστω ότι μας δίνεται ότι η καθυστέρηση στη χρήση του ασανσέρ είναι $Z = |X - Y|$, και καλούμαστε να υπολογίσουμε την μάζα της Z και τη μέση τιμή της. Χρησιμοποιώντας την από κοινού κατανομή, βρίσκουμε:

$$\begin{aligned}
p_Z(0) &= P(Z=0) = P(X=Y) = \frac{1}{6} + \frac{1}{18} + \frac{1}{18} + \frac{1}{18} = \frac{1}{3}, \\
p_Z(1) &= P(Z=1) = P(X=Y+1) + P(X=Y-1) \\
&= \frac{1}{12} + \frac{1}{36} + \frac{1}{36} + \frac{1}{12} + \frac{1}{36} + \frac{1}{36} = \frac{5}{18}, \\
p_Z(2) &= P(Z=2) = P(X=Y+2) + P(X=Y-2) \\
&= \frac{1}{12} + \frac{1}{36} + \frac{1}{12} + \frac{1}{36} = \frac{2}{9}, \\
p_Z(3) &= P(Z=3) = P(X=3,Y=0) + P(X=0,Y=3) \\
&= \frac{1}{12} + \frac{1}{12} = \frac{1}{6},
\end{aligned}$$

και ακολούθως

$$E(|X-Y|) = E(Z) = 0 \times \frac{1}{3} + 1 \times \frac{5}{18} + 2 \times \frac{2}{9} + 3 \times \frac{1}{6} = \frac{11}{9}.$$

Με δεδομένο ότι ο χρήστης είναι στο ισόγειο, ποια είναι η πιθανότητα το ασανσέρ να είναι στον πρώτο όροφο; Έχουμε

$$P(X=1|Y=0) = \frac{P(X=1,Y=0)}{P(Y=0)} = \frac{1/12}{15/36} = \frac{1}{5}.$$

Παράδειγμα 6.6. *(Τελική βαθμολογία)* Βάσει εμπειρικών δεδομένων, ένα καλό μοντέλο για την από κοινού μάζα πιθανότητας του βαθμού $X \in \{0,1,2,\ldots,8\}$ ενός φοιτητή στην τελική εξέταση ενός μαθήματος και του βαθμού $Y \in \{0,1,2\}$ στην πρόοδο αυτού του μαθήματος είναι το ακόλουθο:

x / y	0	1	2	3	4	5	6	7	8	$p_Y(y)$
0	1/18	1/18	1/18	1/27	1/27	1/27	0	0	0	5/18
1	1/18	1/18	1/18	1/27	1/27	1/27	1/18	1/18	1/18	8/18
2	0	0	0	1/27	1/27	1/27	1/18	1/18	1/18	5/18
$p_X(x)$	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9	1/9	

Οι περιθώριες προέκυψαν αθροίζοντας τις αντίστοιχες γραμμές/στήλες. Παρατηρήστε ότι το μοντέλο προβλέπει ότι φοιτητές που λαμβάνουν μεγάλο βαθμό στη μια εξέταση είναι αρκετά πιθανό να λάβουν μεγάλο βαθμό και στην άλλη.

Θα υπολογίσουμε τη μέση τιμή της τελικής βαθμολογίας $Z = X + Y$. Θα υπολογίσουμε καταρχήν την μάζα του Z:

$$\begin{aligned}
p_Z(0) &= p_{XY}(0,0) = \frac{1}{18},\\
p_Z(1) &= p_{XY}(1,0) + p_{XY}(0,1) = \frac{1}{9},\\
p_Z(2) &= p_{XY}(2,0) + p_{XY}(1,1) + p_{XY}(0,2) = \frac{1}{9},\\
p_Z(3) &= p_{XY}(3,0) + p_{XY}(2,1) + p_{XY}(1,2) = \frac{5}{54},\\
p_Z(4) &= p_{XY}(4,0) + p_{XY}(3,1) + p_{XY}(2,2) = \frac{2}{27},\\
p_Z(5) &= p_{XY}(5,0) + p_{XY}(4,1) + p_{XY}(3,2) = \frac{1}{9},
\end{aligned}$$

$$\begin{aligned}
p_Z(6) &= p_{XY}(6,0) + p_{XY}(5,1) + p_{XY}(4,2) = \frac{2}{27},\\
p_Z(7) &= p_{XY}(7,0) + p_{XY}(6,1) + p_{XY}(5,2) = \frac{5}{54},\\
p_Z(8) &= p_{XY}(8,0) + p_{XY}(7,1) + p_{XY}(6,2) = \frac{1}{9},\\
p_Z(9) &= p_{XY}(8,1) + p_{XY}(7,2) = \frac{1}{9},\\
p_Z(10) &= p_{XY}(8,2) = \frac{1}{18}.
\end{aligned}$$

Η μέση τιμή $E(Z)$ προκύπτει

$$\begin{aligned}E(Z) &= 0\times\frac{1}{18}+1\times\frac{1}{9}+2\times\frac{1}{9}+3\times\frac{5}{54}+4\times\frac{2}{27}\\&+5\times\frac{1}{9}+6\times\frac{2}{27}+7\times\frac{5}{54}+8\times\frac{1}{9}+9\times\frac{1}{9}+10\times\frac{1}{18}=5.\end{aligned}$$

Παράδειγμα 6.7. Μερικές φορές το πλήθος των τιμών που παίρνει μια Τ.Μ. Y είναι απαγορευτικά μεγάλο για να αναπαραστήσουμε σε πίνακα την από κοινού μάζα της με μια άλλη Τ.Μ. X.

Έστω, για παράδειγμα, το ακόλουθο πείραμα. Επιλέγουμε στην τύχη έναν από τρεις καλαθοσφαιριστές, και του ζητάμε να εκτελεί ελεύθερες βολές μέχρι να βάλει το πρώτο καλάθι. Οι καλαθοσφαιριστές επιλέγονται χωρίς κάποια προτίμηση ο ένας από τον άλλο, και έχουν ποσοστά επιτυχίας 50%, 33%, και 25%.

Ορίζουμε την Τ.Μ. X ως

$$X=\begin{cases}1/2, & \text{με πιθανότητα } 1/3,\\1/3, & \text{με πιθανότητα } 1/3,\\1/4, & \text{με πιθανότητα } 1/3.\end{cases}$$

Συνεπώς, η X εκφράζει την ευστοχία του τυχαία επιλεγμένου παίκτη. Ορίζουμε επίσης την Y ως μια Γεωμετρική Τ.Μ. με μέση τιμή X. Δηλαδή, η παράμετρος της Τ.Μ. Y είναι επίσης τυχαία!

Η από κοινού μάζα των X,Y μπορεί εύκολα να υπολογιστεί από τον ορισμό της δεσμευμένης πιθανότητας: Για $x\in S_X$ και $y\in S_Y$,

$$p_{XY}(x,y)=P(X=x,Y=y)=P(X=x)P(Y=y|X=x)=\frac{1}{3}(1-x)^{y-1}x,$$

όπου χρησιμοποιήσαμε τον τύπο της μάζας της Γεωμ(x) κατανομής.

Η περιθώρια μάζα της X μας είναι εξ ορισμού γνωστή, ενώ εύκολα μπορούμε να υπολογίσουμε και τις τιμές της περιθώριας μάζας της Y. Για παράδειγμα,

$$\begin{aligned}p_Y(1) &= p_{XY}(1/2,1)+p_{XY}(1/3,1)+p_{XY}(1/4,1)\\&= \frac{1}{3}\left(1-\frac{1}{2}\right)^{1-1}\frac{1}{2}+\frac{1}{3}\left(1-\frac{1}{3}\right)^{1-1}\frac{1}{3}+\frac{1}{3}\left(1-\frac{1}{4}\right)^{1-1}\frac{1}{4}=\frac{13}{36}.\end{aligned}$$

Από τον ορισμό της X γνωρίζουμε πως η πιθανότητα $P(X=1/4)=1/3$. Αλλά δεδομένου ότι το $Y=1$, ποια θα ήταν η πιθανότητα να έχουμε $X=1/4$; Το $X=1/4$ αντιστοιχεί στην περίπτωση που το Y έχει γεωμετρική κατανομή με παράμετρο 1/4, δηλαδή με πιθανότητα «επιτυχίας» που είναι η μικρότερη από τις τρεις δυνατές περιπτώσεις (1/2, 1/3 ή 1/4). Άρα θεωρούμε ότι υπάρχει μικρή πιθανότητα να έχουμε

$X = 1/4$ δεδομένου ότι είχαμε $Y = 1$, δηλαδή «επιτυχία» από το πρώτο κιόλας πείραμα. Με άλλα λόγια, περιμένουμε πως η δεσμευμένη πιθανότητα $P(X = 1/4|Y = 1)$ θα είναι μικρότερη από την αρχική $P(X = 1/4) = 1/3$. Πράγματι, από τον ορισμό της δεσμευμένης πιθανότητας βρίσκουμε

$$P(X = 1/4|Y = 1) = \frac{P(X = 1/4, Y = 1)}{P(Y = 1)} = \frac{p_{XY}(1/4, 1)}{p_Y(1)} = \frac{\frac{1}{3}\left(1-\frac{1}{4}\right)^{1-1}\frac{1}{4}}{13/36} = \frac{3}{13}.$$

Παράδειγμα 6.8. *(Λαχειοφόρος αγορά)* Σε μια λαχειοφόρο αγορά υπάρχουν 10 λαχνοί, εκ των οποίων κερδίζουν οι δύο, από ένα δώρο ο καθένας (τα δύο δώρα είναι πανομοιότυπα). Δύο άτομα αγοράζουν από δύο λαχνούς ο καθένας. Έστω $X, Y \in \{0, 1, 2\}$ το πλήθος των δώρων που κερδίζει ο καθένας.

Θα υπολογίσουμε καταρχήν την από κοινού μάζα των X, Y. Αφού $X, Y \in \{0, 1, 2\}$, πρέπει να υπολογίσουμε 9 τιμές συνολικά. Όμως $p_{XY}(2, 2) = p_{XY}(2, 1) = p_{XY}(1, 2) = 0$, αφού έχουμε μόνο 2 δώρα. Επιπλέον, λόγω συμμετρίας, $p_{XY}(0, 1) = p_{XY}(1, 0)$ και $p_{XY}(0, 2) = p_{XY}(2, 0)$. Άρα, τελικά μας μένει να υπολογίσουμε 4 τιμές της από κοινού πυκνότητας, τις $p_{XY}(0, 0)$, $p_{XY}(1, 1)$, $p_{XY}(1, 0)$, $p_{XY}(2, 0)$.

Η $p_{XY}(0, 0)$ ισούται με την πιθανότητα να πάρουμε 4 λαχνούς από 10, εκ των οποίων δύο κερδίζουν, και να μην επιλέξουμε κανέναν από τους δύο. Άρα,

$$p_{XY}(0, 0) = \frac{\binom{8}{4}}{\binom{10}{4}} = \frac{1}{3}.$$

Η $p_{XY}(2, 0)$ ισούται με την πιθανότητα να επιλέξει ο πρώτος παίκτης 2 λαχνούς ανάμεσα στους 10, και να πετύχει και τους δύο λαχνούς που κερδίζουν. Άρα,

$$p_{XY}(2, 0) = \frac{1}{\binom{10}{2}} = \frac{1}{45}.$$

Η $p_{XY}(1, 0)$ ισούται με την πιθανότητα να επιλέξει ο πρώτος παίκτης 2 λαχνούς ανάμεσα στους 10, και να πετύχει ένα λαχνό που κερδίζει, ενώ ο δεύτερος να επιλέξει 2 λαχνούς ανάμεσα σε 8 λαχνούς που περιέχουν ένα που κερδίζει και να μην τον βρει. Άρα,

$$p_{XY}(1, 0) = \frac{2 \times 8 \times \binom{7}{2}}{\binom{10}{2}\binom{8}{2}} = \frac{4}{15}.$$

Η $p_{XY}(1, 1)$ ισούται με την πιθανότητα να επιλέξει ο πρώτος παίκτης 2 λαχνούς ανάμεσα στους 10, και να πετύχει ένα λαχνό που κερδίζει, ενώ ο δεύτερος να επιλέξει ανάμεσα σε 8 λαχνούς που περιέχουν ένα που κερδίζει και να τον βρει. Άρα,

$$p_{XY}(1, 1) = \frac{2 \times 8 \times 1 \times 7}{\binom{10}{2}\binom{8}{2}} = \frac{4}{45}.$$

Συγκεντρωτικά, έχουμε τον πίνακα

x / y	0	1	2
0	15/45	12/45	1/45
1	12/45	4/45	0
2	1/45	0	0

Βάσει του πίνακα, ποια είναι η πιθανότητα του ενδεχόμενου A να πάρουν και τα δύο δώρα οι δύο διαγωνιζόμενοι; Ποια είναι η πιθανότητα του ενδεχόμενου B ένας (οποιοσδήποτε) από τους δύο να πάρει και τα δύο δώρα; Με χρήση του πίνακα,

$$\begin{aligned} P(A) &= p_{XY}(1,1) + p_{XY}(2,0) + p_{XY}(0,2) = \frac{4}{45} + \frac{1}{45} + \frac{1}{45} = \frac{2}{15}, \\ P(B) &= p_{XY}(2,0) + p_{XY}(0,2) = \frac{1}{45} + \frac{1}{45} = \frac{2}{45}. \end{aligned}$$

Έχοντας την από κοινού μάζα, εύκολα βρίσκουμε πως

$$\begin{aligned} p_X(0) &= p_{XY}(0,0) + p_{XY}(0,1) + p_{XY}(0,2) = \frac{28}{45}, \\ p_X(1) &= p_{XY}(1,0) + p_{XY}(1,1) + p_{XY}(1,2) = \frac{16}{45}, \\ p_X(2) &= p_{XY}(2,0) + p_{XY}(2,1) + p_{XY}(2,2) = \frac{1}{45}, \end{aligned}$$

ενώ λόγω συμμετρίας

$$p_Y(0) = \frac{28}{45}, \quad p_Y(1) = \frac{16}{45}, \quad p_Y(2) = \frac{1}{45},$$

και συγκεντρωτικά έχουμε

x / y	0	1	2	$p_Y(y)$
0	15/45	12/45	1/45	28/45
1	12/45	4/45	0	16/45
2	1/45	0	0	1/45
$p_X(x)$	28/45	16/45	1/45	

Τέλος,

$$E(X) = E(Y) = 0 \times \frac{28}{45} + 1 \times \frac{16}{45} + 2 \times \frac{1}{45} = \frac{2}{5}.$$

6.3 Μέση Τιμή Συνάρτησης Δύο Τυχαίων Μεταβλητών

Λήμμα 6.2. (Μέση τιμή συνάρτησης δύο Τ.Μ.) *Έστω δύο Τ.Μ. X, Y, με σύνολα τιμών S_X, S_Y αντιστοίχως, και από κοινού μάζα πιθανότητας $p_{XY}(x,y)$.*

1. *Έστω Τ.Μ. $Z = g(X,Y)$. Η μέση τιμή της ισούται με*
$$E(Z) = E(g(X,Y)) = \sum_{x\in S_X,\ y\in S_Y} g(x,y)p_{XY}(x,y).$$
2. *$E(aX+bY+c) = aE(X)+bE(Y)+c$, για όλα τα $a,b,c\in\mathbb{R}$.*
3. *Πιο γενικά, αν έχουμε N Τ.Μ. $Z_i = g_i(X,Y)$, $i=1,\ldots,N$, και παραμέτρους $a_1, a_2, \ldots, a_i \in \mathbb{R}$, τότε*
$$E\left(\sum_{i=1}^{N} a_i g_i(X,Y)\right) = \sum_{i=1}^{N} a_i E(g_i(X,Y)).$$

Απόδειξη. 1. Έστω $p_Z(z)$ η μάζα της Z. Παρατηρούμε πως

$$\begin{aligned}
E(Z) &= \sum_{z\in S_Z} z p_Z(z) = \sum_{z\in S_Z}\left(z \sum_{x\in S_X, y\in S_Y : g(x,y)=z} p_{XY}(x,y)\right)\\
&= \sum_{z\in S_Z}\left(\sum_{x\in S_X, y\in S_Y : g(x,y)=z} g(x,y)p_{XY}(x,y)\right)\\
&= \sum_{x\in S_X,\ y\in S_Y} g(x,y)p_{XY}(x,y).
\end{aligned}$$

Η πρώτη ισότητα ισχύει εξ ορισμού. Στην δεύτερη ισότητα χρησιμοποιήσαμε τον νόμο της ολικής πιθανότητας, όπως αυτός εμφανίζεται στο δεύτερο σκέλος του Λήμματος 6.1.

2. Παρατηρούμε πως:

$$\begin{aligned}
E(aX+bY+c) &= \sum_{x\in S_X,\ y\in S_Y} (ax+by+c)p_{XY}(x,y)\\
&= a\sum_{x\in S_X,\ y\in S_Y} x p_{XY}(x,y) + b\sum_{x\in S_X,\ y\in S_Y} y p_{XY}(x,y)\\
&\quad + c\sum_{x\in S_X,\ y\in S_Y} p_{XY}(x,y) = aE(X)+bE(Y)+c,
\end{aligned}$$

όπου στην τρίτη ισότητα χρησιμοποιήσαμε το προηγούμενο σκέλος.

3.

$$E\left(\sum_{i=1}^{N} a_i g_i(X,Y)\right) = \sum_{x\in S_X,\ y\in S_Y} \left(\sum_{i=1}^{N} a_i g_i(X,Y)\right) p_{XY}(x,y)$$
$$= \sum_{i=1}^{N} a_i \sum_{x\in S_X,\ y\in S_Y} g_i(X,Y) p_{XY}(x,y) = \sum_{i=1}^{N} a_i E(g_i(X,Y)).$$

□

Παρατήρηση: Όπως είδαμε και σε προηγούμενα παραδείγματα, πολύ συχνά χρειάζεται να υπολογίσουμε την μέση τιμή Τ.Μ. Z που είναι συναρτήσεις δύο άλλων Τ.Μ. Σε πολλές περιπτώσεις, το άνω λήμμα απλουστεύει σημαντικά τους υπολογισμούς που χρειάζονται, γιατί μας επιτρέπει να αποφύγουμε τον υπολογισμό της μάζας της Z. Δείτε τα παραδείγματα που ακολουθούν.

Παράδειγμα 6.9. *(Ασανσέρ — συνέχεια)* Σε συνέχεια του Παραδείγματος 6.5, μπορούμε να υπολογίσουμε το μέσο χρόνο αναμονής $E(Z) = E(|X-Y|)$ και ως εξής:

$$\begin{aligned} E(|X-Y|) &= |0-0|\times\frac{1}{6} + (|1-1|+|2-2|+|3-3|)\times\frac{1}{18} \\ &\quad +(|1-0|+|2-0|+|3-0|+|0-1|+|0-2|+|0-3|)\times\frac{1}{12} \\ &\quad +(|2-1|+|3-1|+|1-2|+|3-2|+|1-3|+|2-3|)\times\frac{1}{36} \\ &= \frac{11}{9}. \end{aligned}$$

Παράδειγμα 6.10. *(Τελική Βαθμολογία — συνέχεια)* Σε συνέχεια του Παραδείγματος 6.6,

$$\begin{aligned} E(X) &= (0+1+2+\cdots+8)\times\frac{1}{9} = 4, \\ E(Y) &= 0\times\frac{5}{18} + 1\times\frac{8}{18} + 2\times\frac{5}{18} = 1, \\ E(X+Y) &= E(X)+E(Y) = 5. \end{aligned}$$

Παράδειγμα 6.11. *(Δύο ζάρια — συνέχεια)* Έστω το ακόλουθο παιχνίδι: ρίχνουμε δύο δίκαια και ανεξάρτητα ζάρια, και κερδίζουμε $2\max\{X,Y\}+3\min\{X,Y\}$ ευρώ. Για να παίξουμε όμως, πρέπει αρχικά να καταβάλουμε 15 ευρώ. Έστω $P = 2\max\{X,Y\}+$

$3\min\{X,Y\}-15$ το καθαρό κέρδος. Μας συμφέρει να παίξουμε; Για να απαντήσουμε το ερώτημα, πρέπει να υπολογίσουμε την μέση τιμή

$$E(P) = E(2\max\{X,Y\} + 3\min\{X,Y\} - 15).$$

Με δεδομένο ότι γνωρίζουμε τις μέσες τιμές $E(\max\{X,Y\})$ και $E(\min\{X,Y\})$ από προηγούμενα παραδείγματα, μας συμφέρει να χρησιμοποιήσουμε το τρίτο σκέλος του άνω λήμματος ως εξής:

$$\begin{aligned} E(2\max\{X,Y\} + 3\min\{X,Y\} - 15) &= 2E(\max\{X,Y\}) + 3E(\min\{X,Y\}) - 15 \\ &= 2 \times \frac{161}{36} + 3 \times \frac{91}{36} - 15 = \frac{55}{36}. \end{aligned}$$

Επειδή η μέση τιμή προκύπτει θετική, μας συμφέρει να παίξουμε.

Εναλλακτικά, χωρίς χρήση του λήμματος, θα έπρεπε να υπολογίσουμε την μάζα του κέρδους P, και από αυτή την μέση τιμή της P με χρήση του ορισμού της μέσης τιμής. Ο υπολογισμός αυτός είναι χρονοβόρος.

Ορισμός 6.2. *Η <u>συνδιακύμανση</u>* $\mathrm{COV}(X,Y)$ *μεταξύ δύο διακριτών Τ.Μ.* X,Y *ορίζεται ως:*

$$\mathrm{COV}(X,Y) \triangleq E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big].$$

Παρατήρηση: Μια πρώτη διαισθητική ερμηνεία της συνδιακύμανσης είναι πως, όταν $\mathrm{COV}(X,Y) > 0$, τότε οι δύο Τ.Μ. X,Y τείνουν να παίρνουν τις «μεγάλες» και τις «μικρές» τιμές τους *ταυτόχρονα*. Αντίστοιχα, όταν $\mathrm{COV}(X,Y) < 0$, τότε όταν η μία Τ.Μ. παίρνει μεγάλες τιμές η άλλη τείνει να παίρνει μικρές τιμές. Άρα η συνδιακύμανση $\mathrm{COV}(X,Y)$ παρέχει μια πρώτη ένδειξη για τη σχέση ανάμεσα στις X,Y. Δείτε τις συνδιακυμάνσεις των επόμενων παραδειγμάτων.

Παράδειγμα 6.12. *(Δύο προγράμματα — συνέχεια)* Για τις Τ.Μ. X,Y του Παραδείγματος 6.1 έχουμε

$$E(X) = E(Y) = 0 \times \frac{4}{9} + 1 \times \frac{4}{9} + 2 \times \frac{1}{9} = \frac{2}{3}.$$

Από τον ορισμό της συνδιακύμανσης και την από κοινού μάζα των X,Y, εύκολα υπολογίζουμε

$$\begin{aligned} \mathrm{COV}(X,Y) &= \frac{1}{9}\Big(0-\frac{2}{3}\Big)\Big(0-\frac{2}{3}\Big) + \frac{2}{9}\Big(1-\frac{2}{3}\Big)\Big(0-\frac{2}{3}\Big) + \frac{1}{9}\Big(2-\frac{2}{3}\Big)\Big(0-\frac{2}{3}\Big) \\ &\quad + \frac{2}{9}\Big(0-\frac{2}{3}\Big)\Big(1-\frac{2}{3}\Big) + \frac{2}{9}\Big(1-\frac{2}{3}\Big)\Big(1-\frac{2}{3}\Big) + \frac{1}{9}\Big(0-\frac{2}{3}\Big)\Big(2-\frac{2}{3}\Big) \\ &= -\frac{2}{9}, \end{aligned}$$

όπου στον υπολογισμό παραλείψαμε τα τρία ζεύγη τιμών (x,y) με μηδενική πιθανότητα.

Λήμμα 6.3. (Ιδιότητες συνδιακύμανσης) *Έστω X, Y διακριτές Τ.Μ.*

1. $\mathrm{COV}(X, X) = \mathrm{VAR}(X)$.
2. $\mathrm{COV}(X, -X) = -\mathrm{VAR}(X)$.
3. $\mathrm{COV}(X, Y) = E(XY) - E(X)E(Y)$.
4. $\mathrm{VAR}(X + Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\mathrm{COV}(X, Y)$.

Απόδειξη. 1. Από τον ορισμό της συνδιακύμανσης, άμεσα προκύπτει πως

$$\mathrm{COV}(X, X) = E\big[\big(X - E(X)\big)\big(X - E(X)\big)\big] = E\big[(X - E(X))^2\big] = \mathrm{VAR}(X).$$

2. Ομοίως:

$$\begin{aligned}\mathrm{COV}(X, -X) = E\big[\big(X - E(X)\big)\big(-X + E(X)\big)\big] \\ = E\big[-(X - E(X))^2\big] = -\mathrm{VAR}(X).\end{aligned}$$

3. Από τον ορισμό της συνδιακύμανσης, άμεσα έχουμε:

$$\begin{aligned}\mathrm{COV}(X, Y) &= E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big] \\ &= E\big[XY - E(X)Y - XE(Y) + E(X)E(Y)\big] \\ &= E[XY] - E[E(X)Y] - E[XE(Y)] + E[E(X)E(Y)]] \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y).\end{aligned}$$

4. Ξεκινώντας από τον ορισμό της διασποράς, έχουμε:

$$\begin{aligned}&\mathrm{VAR}(X + Y) \\ &= E\left[((X + Y) - E(X + Y))^2\right] = E\left[((X - E(X)) + (Y - E(Y)))^2\right] \\ &= E\left[(X - E(X))^2\right] + E\left[(Y - E(Y))^2\right] + 2E\left[(X - E(Y))(Y - E(Y))\right] \\ &= \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\mathrm{COV}(X, Y).\end{aligned}$$

□

Παρατήρηση: Το τρίτο σκέλος του άνω λήμματος συχνά μας επιτρέπει να υπολογίσουμε την συνδιακύμανση με λιγότερες πράξεις. Στα επόμενα παραδείγματα, βεβαιωθείτε ότι η χρήση του ορισμού θα έδινε το ίδιο αποτέλεσμα.

Παράδειγμα 6.13. *(Δύο ζάρια — συνέχεια)* Έστω πως ρίχνουμε δύο δίκαια και ανεξάρτητα ζάρια X και Y, έτσι ώστε η από κοινού τους μάζα αν είναι η πρώτη του Παραδείγματος 6.2. Από το Παράδειγμα 4.4 έχουμε $E(X) = E(Y) = \frac{7}{2}$. Άρα:

$$\begin{aligned} E(XY) &= (1+2+3+4+5+6)\times(1+2+3+4+5+6)\times\frac{1}{36} = \frac{49}{4}, \\ \mathrm{COV}(X,Y) &= E(XY) - E(X)E(Y) = \frac{49}{4} - \frac{7}{2}\times\frac{7}{2} = 0. \end{aligned}$$

Αν όμως τα ζάρια έχουν από κοινού μάζα την δεύτερη του Παραδείγματος 6.2, τότε, αφού και πάλι $E(X) = E(Y) = \frac{7}{2}$, έχουμε:

$$\begin{aligned} E(XY) &= (1^2+2^2+3^3+4^2+5^2+6^2)\times\frac{1}{6} = \frac{91}{6}, \\ \mathrm{COV}(X,Y) &= E(XY) - E(X)E(Y) = \frac{91}{6} - \frac{7}{2}\times\frac{7}{2} = \frac{35}{12}. \end{aligned}$$

Σε συνέχεια του Παραδείγματος 6.3, θα υπολογίσουμε επίσης την συνδιακύμανση των $W = \max\{X, Y\}, Z = \min\{X, Y\}$ όταν τα ζάρια είναι δίκαια και ανεξάρτητα. Από το Παράδειγμα 4.4 έχουμε $E(W) = 161/36$ και $E(Z) = 91/36$. Συνεπώς:

$$\begin{aligned} E(WZ) &= (1^2+2^2+3^2+4^2+5^2+6^2)\times\frac{1}{36} \\ &\quad +(1\times(2+3+4+5+6)+2\times(3+4+5+6) \\ &\quad +3\times(4+5+6)+4\times(5+6)+5\times 6)\times\frac{2}{36} = \frac{49}{4}, \\ \mathrm{COV}(W,Z) &= E(WZ) - E(W)E(Z) = \frac{49}{4} - \frac{161}{36}\times\frac{91}{36} = \frac{1225}{1296}. \end{aligned}$$

Παράδειγμα 6.14. *(Τελική βαθμολογία — συνέχεια)* Σε συνέχεια των Παραδειγμάτων 6.6 και 6.10, η συνδιακύμανση των βαθμών X, Y υπολογίζεται ως εξής:

$$\begin{aligned} E(XY) &= [1\times(1+2+6+7+8)+2\times(6+7+8)]\times\frac{1}{18} \\ &\quad +(1\times(3+4+5)+2\times(3+4+5))\times\frac{1}{27} = 5, \\ \mathrm{COV}(X,Y) &= E(XY) - E(X)E(Y) = 5 - 4\times 1 = 1. \end{aligned}$$

Παράδειγμα 6.15. *(Ασανσέρ — συνέχεια)* Σε συνέχεια του Παραδείγματος 6.5,

$$\begin{aligned} E(X) = E(Y) &= 0\times\frac{15}{36} + (1+2+3)\times\frac{7}{36} = \frac{7}{6}, \\ E(XY) &= (1^2+2^2+3^3)\times\frac{1}{18} \\ &\quad +(1+2)\times(3+4+5)\times\frac{1}{36} = \frac{25}{18}, \\ \mathrm{COV}(X,Y) &= E(XY) - E(X)E(Y) = \frac{25}{18} - \frac{7}{6}\times\frac{7}{6} = \frac{1}{36}. \end{aligned}$$

Παράδειγμα 6.16. *(Συντελεστής συσχέτισης)* Ο συντελεστής συσχέτισης δύο τυχαίων μεταβλητών X, Y, ορίζεται ως

$$\rho_{X,Y} = \frac{\text{COV}(X,Y)}{[\text{VAR}(X)\text{VAR}(Y)]^{\frac{1}{2}}}.$$

Θα δείξουμε ότι αν $Y = aX + b$, όπου η σταθερά $a \neq 0$, τότε ο συντελεστής συσχέτισης $\rho_{X,Y}$ των X, Y, ισούται με $+1$ ή με -1.

Καταρχάς παρατηρούμε ότι

$$\begin{aligned} \text{COV}(X,Y) &= E(XY) - E(X)E(Y) = E(X(aX+b)) - E(X)E(aX+b) \\ &= E(aX^2 + bX) - E(X)(aE(X) + b) \\ &= aE(X^2) + bE(X) - aE(X^2) - bE(X) \\ &= a\text{VAR}(X), \end{aligned}$$

Επίσης, έχουμε

$$\text{VAR}(Y) = \text{VAR}(aX + b) = a^2\text{VAR}(X),$$

συνεπώς,

$$\rho_{X,Y} = \frac{\text{COV}(X,Y)}{[\text{VAR}(X)\text{VAR}(Y)]^{\frac{1}{2}}} = \frac{a\text{VAR}(X)}{[\text{VAR}(X)a^2\text{VAR}(X)]^{\frac{1}{2}}} = \frac{a}{|a|} = \begin{cases} 1, & a > 0, \\ -1, & a < 0. \end{cases}$$

Μάλιστα, μπορούμε να δείξουμε ότι πάντα $|\rho_{X,Y}| \leq 1$, επομένως κατά μία έννοια οι άνω περιπτώσεις είναι ακραίες.

6.4 Ανεξάρτητες Τυχαίες Μεταβλητές

Ορισμός 6.3. (Ζεύγη ανεξάρτητων διακριτών Τ.Μ.) *Δύο διακριτές Τ.Μ. X, Y, με σύνολα τιμών S_X, S_Y αντιστοίχως, καλούνται ανεξάρτητες αν για οποιαδήποτε υποσύνολα $A \subseteq S_X$, $B \subseteq S_Y$, ισχύει*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B). \tag{6.1}$$

Παρατηρήσεις

1. Έχουμε ορίσει ήδη την ανεξαρτησία ενδεχόμενων και την ανεξαρτησία υποπειραμάτων (που βασίζεται στην ανεξαρτησία ενδεχόμενων).

2. Με τον άνω ορισμό εισάγουμε και την ανεξαρτησία τυχαίων μεταβλητών. Και αυτή η ανεξαρτησία βασίζεται στην ανεξαρτησία ενδεχόμενων. Πράγματι, δύο Τ.Μ. είναι ανεξάρτητες αν δύο οποιαδήποτε ενδεχόμενα που αφορούν το καθένα αποκλειστικά τη μία από τις δύο Τ.Μ. είναι ανεξάρτητα.

Λήμμα 6.4. (Κριτήριο ανεξαρτησίας διακριτών Τ.Μ.) *Δύο διακριτές Τ. Μ. X, Y, με σύνολα τιμών S_X, S_Y, από κοινού μάζα $p_{XY}(x, y)$, και περιθώριες μάζες $p_X(x)$, $p_Y(y)$, είναι ανεξάρτητες αν και μόνο αν*

$$p_{XY}(x, y) = p_X(x)p_Y(y), \quad \forall x \in S_X,\ \forall y \in S_Y. \tag{6.2}$$

Απόδειξη. Έστω καταρχήν πως οι X, Y είναι ανεξάρτητες. Τότε η εξίσωση (6.2) προκύπτει, για κάθε x, y, με εφαρμογή του ορισμού της ανεξαρτησίας για τα σύνολα $A = \{x\}$, $B = \{y\}$. Αντιστρόφως, έστω πως ισχύει η (6.2). Έστω δύο υποσύνολα $A \subseteq S_X$, $B \subseteq S_Y$. Θα δείξουμε ότι ισχύει η (6.1). Πράγματι:

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x \in A,\ y \in B} p_{XY}(x, y) = \sum_{x \in A} \sum_{y \in B} p_X(x)p_Y(y) \\ &= \sum_{x \in A} \Big(p_X(x) \sum_{y \in B} p_Y(y) \Big) = \Big(\sum_{x \in A} p_X(x) \Big) \times \Big(\sum_{y \in B} p_Y(y) \Big) \\ &= P(X \in A)P(Y \in B), \end{aligned}$$

και η απόδειξη ολοκληρώθηκε. (Στην πρώτη ισότητα χρησιμοποιήσαμε το Λήμμα 6.1 ενώ στην δεύτερη την υπόθεση.) □

Παρατηρήσεις

1. Παρατηρήστε πως όταν οι X, Y είναι ανεξάρτητες, τότε αν $P(X=x)=p_X(x)>0$ θα έχουμε

$$\begin{aligned} P(Y=y|X=x) &= \frac{P(X=x,Y=y)}{P(X=x)} = \frac{p_{XY}(x,y)}{p_X(x)} \\ &= \frac{p_X(x)p_Y(y)}{p_X(x)} = p_Y(y) = P(Y=y). \end{aligned}$$

   Μια ανάλογη απόδειξη θα ισχύει αν αλλάξουμε τις θέσεις των X, Y. Άρα τελικά, αν τα X, Y είναι ανεξάρτητα, θα ισχύει ότι:

$$\begin{aligned} \forall x \in S_X, \quad P(X=x)>0 &\Rightarrow P(Y=y|X=x)=P(Y=y), \\ \forall y \in S_Y, \quad P(Y=y)>0 &\Rightarrow P(X=x|Y=y)=P(X=x). \end{aligned}$$

   Μπορούμε να αποδείξουμε, επιπλέον, ότι αν ισχύουν οι άνω, τότε οι X, Y είναι ανεξάρτητες.

2. Το αναγκαίο και ικανό κριτήριο του Λήμματος 6.4 είναι το βασικό εργαλείο που χρησιμοποιούμε για να διαπιστώνουμε την ανεξαρτησία δύο διακριτών Τ.Μ. Δείτε το επόμενο παράδειγμα.

Παράδειγμα 6.17. *(Δύο ζάρια — συνέχεια)* Παρατηρήστε πως όταν η από κοινού μάζα των αποτελεσμάτων X, Y δύο ζαριών είναι η πρώτη από τις δύο του Παραδείγματος 6.2, τότε σύμφωνα με το άνω κριτήριο οι X, Y είναι ανεξάρτητες, γιατί για κάθε $x, y \in \{1, \ldots, 6\}$ θα έχουμε

$$p_{XY}(x,y) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = p_X(x)p_Y(y).$$

Αν όμως η από κοινού μάζα των X, Y είναι η δεύτερη του Παραδείγματος 6.2, τότε οι X, Y δεν είναι ανεξάρτητες, γιατί, για παράδειγμα

$$p_{XY}(1,2) = 0 \neq \frac{1}{6} \times \frac{1}{6} = p_X(1)p_Y(2).$$

Παρατηρήσεις

1. Σε ποια άλλα από τα προηγούμενα παραδείγματα είχαμε ανεξάρτητες Τ.Μ.; Σε όλες τις περιπτώσεις, βεβαιωθείτε ότι συμβαδίζει η διαίσθησή σας με το κριτήριο του Λήμματος 6.4.

2. Συχνά η ανεξαρτησία δύο Τ.Μ. δεν είναι το ζητούμενο, αλλά δίνεται ως υπόθεση. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 6.18. *(Ελάχιστο δύο ανεξάρτητων γεωμετρικών Τ.Μ.)* Έστω X και Y δύο ανεξάρτητες γεωμετρικές Τ.Μ., με παραμέτρους p_1 και p_2, αντίστοιχα. Θα υπολογίσουμε τη μάζα της νέας Τ.Μ. $Z = \min(X, Y)$.

Παρατηρούμε ότι το ελάχιστο Z είναι τουλάχιστον k, αν και μόνο αν και οι δύο Τ.Μ. X, Y παίρνουν τιμές τουλάχιστον k. Οπότε, για κάθε $k \geq 1$,

$$\begin{aligned} P(Z \geq k) &= P(\min(X,Y) \geq k) = P(X \geq k,\ Y \geq k) = P(X \geq k)P(Y \geq k) \\ &= (1-p_1)^{k-1}(1-p_2)^{k-1} = [(1-p_1)(1-p_2)]^{k-1}, \end{aligned}$$

όπου η τρίτη ισότητα ισχύει λόγω ανεξαρτησίας και η τέταρτη είναι γνωστή ιδιότητα της γεωμετρικής κατανομής.

Αν τώρα ορίσουμε $q = 1 - (1-p_1)(1-p_2)$, εφόσον

$$P(Z \geq k) = P(Z = k) + P(Z \geq k+1),$$

έχουμε

$$P(Z = k) = P(Z \geq k) - P(Z \geq k+1) = (1-q)^{k-1} - (1-q)^k = q(1-q)^{k-1}.$$

Συνεπώς, η Z ακολουθεί επίσης γεωμετρική κατανομή, αλλά με παράμετρο $q = 1-(1-p_1)(1-p_2)$.

Αυτό εξηγείται εύκολα: Αν διεξάγουμε ταυτόχρονα δύο ακολουθίες ανεξάρτητων πειραμάτων, μπορούμε να ορίσουμε ένα ολικό πείραμα στο οποίο θα έχουμε επιτυχία την πρώτη φορά που θα έχει επιτυχία ένα από τα δύο πειράματα. Σ' αυτή την περίπτωση, το πλήθος των επαναλήψεων που απαιτούνται για την πρώτη επιτυχία του ολικού πειράματος είναι $Z = \min(X, Y)$ όπου X, Y είναι το πλήθος επαναλήψεων μέχρι την πρώτη επιτυχία στο πρώτο και το δεύτερο πείραμα, αντίστοιχα. Το ολικό πείραμα αποτυγχάνει όταν αποτύχουν και τα δύο επί μέρους, δηλαδή με πιθανότητα $(1-p_1)(1-p_2)$, άρα η πιθανότητα επιτυχίας είναι $1-(1-p_1)(1-p_2) = q$. Συνεπώς, το πλήθος των προσπαθειών του ολικού πειράματος μέχρι την πρώτη επιτυχία ακολουθεί γεωμετρική κατανομή με παράμετρο q.

Λήμμα 6.5. (Ιδιότητες ανεξάρτητων διακριτών Τ.Μ.) *Έστω X, Y ανεξάρτητες διακριτές Τ.Μ.*

1. *Έστω συναρτήσεις $g : S_X \to \mathbb{R}$, και $h : S_Y \to \mathbb{R}$. Θα ισχύει*

$$E(g(X)h(Y)) = E(g(X))E(h(Y))).$$

Ειδική περίπτωση της άνω είναι η

$$E(XY) = E(X)E(Y).$$

2. $\mathrm{COV}(X, Y) = 0$.

3. $\mathrm{VAR}(X+Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y)$.

Απόδειξη. 1. Εφαρμόζοντας το Λήμμα 6.2 έχουμε:

$$\begin{aligned} E(g(X)h(Y)) &= \sum_{x\in S_X,\ y\in S_Y} h(x)g(y)p_{XY}(x,y) \\ &= \sum_{x\in S_X}\sum_{y\in S_Y} g(x)h(y)p_X(x)p_Y(y) \\ &= \sum_{x\in S_X}\Big(g(x)p_X(x)\sum_{y\in S_Y} h(y)p_Y(y)\Big) \\ &= \Big(\sum_{x\in S_X} g(x)p_X(x)\Big)\Big(\sum_{y\in S_Y} h(y)p_Y(y)\Big) = E(g(X))E(h(Y)). \end{aligned}$$

2. Προκύπτει άμεσα από το πρώτο σκέλος και την εφαρμογή του Λήμματος 6.3.

3. Προκύπτει άμεσα από το δεύτερο σκέλος και την εφαρμογή του Λήμματος 6.3.

□

Παράδειγμα 6.19. *(Δύο ζάρια — συνέχεια)* Έστω πως η από κοινού μάζα των αποτελεσμάτων X, Y δύο ζαριών είναι η πρώτη από τις δύο του Παραδείγματος 6.2, οπότε οι X, Y είναι ανεξάρτητες. Θα έχουμε, για το άθροισμά τους $X+Y$:

$$\text{VAR}(X+Y) = \text{VAR}(X) + \text{VAR}(Y) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}.$$

Παρατηρήστε επίσης ότι

$$E(X+Y) = E(X) + E(Y) = \frac{7}{2} + \frac{7}{2} = 7.$$

Πρέπει να τονιστεί ότι η τελευταία ισότητα θα ίσχυε ακόμα και αν τα X, Y ΔΕΝ ήταν ανεξάρτητα, καθώς το Λήμμα 6.2 (πάνω στο οποίο βασιστήκαμε) δεν απαιτεί ανεξαρτησία!

Παράδειγμα 6.20. Θα δείξουμε ένα παράδειγμα δύο διακριτών Τ.Μ. X, Y με διασπορές $\text{VAR}(X) = \text{VAR}(Y) > 0$, αλλά τέτοιες ώστε να έχουμε $\text{VAR}(X+Y) = 0$.

Έστω $X \sim \text{Bern}(1/2)$ οπότε $\text{VAR}(X) = (1/2)(1/2) = 1/4 > 0$, και έστω $Y = -X$ οπότε $\text{VAR}(Y) = \text{VAR}((-1)X) = (-1)^2\text{VAR}(X) = 1/4 > 0$. Αλλά εφόσον εξ ορισμού η $X+Y=0$ είναι απλώς μια σταθερά, έχουμε $\text{VAR}(X+Y)=0$, το οποίο φυσικά δεν ισούται με τον άθροισμα των διασπορών $\text{VAR}(X)+\text{VAR}(Y) = 1/2$, όπως προβλέπει το άνω λήμμα. Αυτό συμβαίνει γιατί προφανώς οι X, Y δεν είναι ανεξάρτητες.

Παράδειγμα 6.21. *(Μη ανεξάρτητες Τ.Μ. με μηδενική συνδιακύμανση)* Είναι δυνατόν δύο Τ.Μ. να μην είναι ανεξάρτητες, να έχουν όμως μηδενική συνδιακύμανση. Για παράδειγμα, έστω δύο διακριτές Τ.Μ. X, Y με από κοινού μάζα p_{XY} όπως στον πιο κάτω πίνακα:

x / y	-1	0	1	$p_Y(y)$
-1	0	1/4	0	1/4
0	1/4	0	1/4	1/2
1	0	1/4	0	1/4
$p_X(x)$	1/4	1/2	1/4	

Οι περιθώριες μάζες $p_X(x), p_Y(y)$ έχουν επίσης υπολογιστεί στον πίνακα. Οι X, Y δεν είναι ανεξάρτητες, αφού

$$p_{XY}(0,0) = 0 \neq \frac{1}{2} \times \frac{1}{2} = p_X(0)p_Y(0).$$

Για τη συνδιακύμανσή τους, πρώτα υπολογίζουμε

$$E(X) = (-1) \times \frac{1}{4} + 0 \times \frac{1}{2} + 1 \times \frac{1}{4} = 0,$$

και παρομοίως $E(Y) = 0$ αφού οι X και Y έχουν την ίδια μάζα. Επιπλέον παρατηρούμε πως

$$E(XY) = \sum_{x=-1,0,1} \sum_{y=-1,0,1} xy p_{XY}(x,y) = 0,$$

διότι *όλοι* οι όροι του πιο πάνω αθροίσματος είναι μηδενικοί! (Ή $x = 0$ ή $y = 0$ ή $p_{XY}(x,y) = 0$, για όλα τα δυνατά ζεύγη τιμών (x,y).) Άρα,

$$\text{COV}(X,Y) = E(XY) - E(X)E(Y) = 0 - 0 \times 0 = 0.$$

Ορισμός 6.4. (Ασυσχέτιστες Τ.Μ.) *Δύο Τ.Μ. X, Y, καλούνται ασυσχέτιστες αν η συνδιακύμανσή τους $\text{COV}(X,Y)$ είναι μηδενική, δηλαδή $\text{COV}(X,Y) = 0$. Από τα προηγούμενα είδαμε ότι δύο Τ.Μ. που είναι ανεξάρτητες είναι και ασυσχέτιστες, αλλά το αντίστροφο μπορεί να μην ισχύει.*

6.5 Άθροισμα Ανεξάρτητων Τυχαίων Μεταβλητών

Παράδειγμα 6.22. *(Συνέλιξη)* Έστω δύο ανεξάρτητες διακριτές Τ.Μ. X, Y με $S_X = S_Y = \mathbb{Z}$ και μάζες $p_X(x)$, $p_Y(y)$ αντίστοιχα. Θα δείξουμε ότι η νέα Τ.Μ. $X+Y$ έχει μάζα:

$$p_{X+Y}(m) = \sum_{k=-\infty}^{\infty} p_X(k) p_Y(m-k), \quad m \in \mathbb{Z}.$$

Η πιο πάνω έκφραση είναι γνωστή ως η *συνέλιξη* των δύο μαζών $p_X(x)$ και $p_Y(y)$, και εμφανίζεται συχνά σε εφαρμογές, ακόμα και εκτός θεωρίας πιθανοτήτων. Αργότερα θα δούμε και την συνεχή εκδοχή της, που έχει τη μορφή ολοκληρώματος αντί για αθροίσματος, και της οποίας η χρήση είναι ακόμα πιο διαδεδομένη.

Πράγματι, παρατηρούμε πως, για οποιοδήποτε τιμή $m \in \mathbb{Z}$, η $X+Y$ ισούται με m αν και μόνο αν η $X = k$ και η $Y = m-k$ για κάποιο $k \in \mathbb{Z}$. Συνεπώς:

$$\begin{aligned} p_{X+Y}(m) &= P(X+Y=m) = P\left(\bigcup_{k=-\infty}^{\infty} (\{X=k\} \cap \{Y=m-k\}) \right) \\ &= \sum_{k=-\infty}^{\infty} P(\{X=k\} \cap \{Y=m-k\}) = \sum_{k=-\infty}^{\infty} P(X=k)P(Y=m-k) \\ &= \sum_{k=-\infty}^{\infty} p_X(k) p_Y(m-k). \end{aligned}$$

Η πρώτη και η τελευταία ισότητα προκύπτουν από τον ορισμό της μάζας, η δεύτερη προκύπτει γράφοντας το ενδεχόμενο $\{X+Y=m\}$ σαν ένωση ξένων ενδεχόμενων, η τρίτη προκύπτει ακριβώς επειδή τα ενδεχόμενα είναι ξένα, ενώ η τέταρτη λόγω της ανεξαρτησίας των X, Y.

Παρατήρηση: Παρατηρήστε ότι αν μια Τ.Μ. X έχει σύνολο τιμών $S_X \subset \mathbb{Z}$ μπορούμε να θέσουμε ως νέο σύνολο τιμών της το $\mathbb{Z}$, θέτοντας μηδενικές τις πιθανότητες των ακεραίων εκτός του αρχικού S_X. Αντιστρόφως, μπορούμε να αφαιρέσουμε από το σύνολο τιμών S_X οποιαδήποτε τιμή x για την οποία $p_X(x) = 0$. Και οι δύο αυτές διαδικασίες δεν αλλάζουν ουσιωδώς την X, και γενικεύουν την εφαρμογή του άνω παραδείγματος. Δείτε τα παραδείγματα που ακολουθούν.

Παράδειγμα 6.23. *(Άθροισμα ανεξάρτητων Poisson Τ.Μ.)* Αν οι Τ.Μ. X και Y είναι ανεξάρτητες, με κατανομές $\text{Poisson}(\lambda)$ και $\text{Poisson}(\mu)$ αντίστοιχα, θα δείξουμε ότι η Τ.Μ. $X+Y$ έχει κατανομή $\text{Poisson}(\lambda+\mu)$.

Πράγματι, έστω $p_X(x)$ και $p_Y(y)$ οι μάζες των X και Y, αντίστοιχα, και έστω $p_{X+Y}(m)$ η μάζα της $X+Y$. Από το προηγούμενο παράδειγμα, έχουμε:

$$\begin{aligned} p_{X+Y}(m) &= \sum_{k=-\infty}^{\infty} p_X(k)\, p_Y(m-k) = \sum_{k=0}^{m} \Big(\frac{e^{-\lambda}\lambda^k}{k!}\Big)\Big(\frac{e^{-\mu}\mu^{m-k}}{(m-k)!}\Big) \\ &= \frac{e^{-(\lambda+\mu)}}{m!}\sum_{k=0}^{m}\frac{m!}{k!(m-k)!}\lambda^k\mu^{m-k} = \frac{e^{-(\lambda+\mu)}}{m!}(\lambda+\mu)^m. \end{aligned}$$

Παρατηρήστε ότι σε αυτή την περίπτωση και οι δύο Τ.Μ. που απαρτίζουν το άθροισμα είναι μη αρνητικές, και συνεπώς δεν προκύπτει άπειρο άθροισμα όπως στην γενική περίπτωση του προηγούμενου παραδείγματος. Επίσης, η τελευταία ισότητα προκύπτει από το διωνυμικό θεώρημα:

$$(\alpha+b)^n = \sum_{k=0}^{n}\binom{n}{k}\alpha^k b^{n-k}.$$

Συνεπώς, $X+Y \sim \text{Poisson}(\lambda+\mu)$.

Παράδειγμα 6.24. *(Άθροισμα ανεξάρτητων διωνυμικών Τ.Μ.)* Σε αυτό το παράδειγμα θα δείξουμε ότι το άθροισμα δύο διωνυμικών Τ.Μ. X, Y με παραμέτρους n_1, p και n_2, p αντίστοιχα, ανεξάρτητων μεταξύ τους, έχει Διων(n_1+n_2, p) κατανομή.

Έστω $p_X(x)$ και $p_Y(y)$ οι μάζες των X και Y, αντίστοιχα. Παρατηρήστε πως

$$\begin{aligned} p_{X+Y}(m) &= \sum_{k=-\infty}^{\infty} p_X(k)p_Y(m-k) \\ &= \sum_{k=\max\{0,m-n_2\}}^{\min\{n_1,m\}} \binom{n_1}{k}p^k(1-p)^{n_1-k}\binom{n_2}{m-k}p^{m-k}(1-p)^{n_2-m+k} \\ &= (p^m(1-p)^{n_1+n_2-m}) \sum_{k=\max\{0,m-n_2\}}^{\min\{n_1,m\}}\binom{n_1}{k}\binom{n_2}{m-k}. \end{aligned} \tag{6.3}$$

Τα όρια στο άθροισμα προέκυψαν κρατώντας μόνο τις τιμές του k για τις οποίες είναι θετικές και οι δύο μάζες. Παρατηρήστε ότι όταν $m > n_1+n_2$ ή $m < 0$, το άνω άθροισμα είναι κενό και προκύπτει πως $p_{X+Y}(m) = 0$, όπως και με την Διων(n_1+n_2, p) κατανομή. Αρκεί λοιπόν να επικεντρωθούμε στην περίπτωση $0 \le m \le n_1+n_2$.

Σε αυτή την περίπτωση, παρατηρήστε πως

$$\sum_{k=\max\{0,m-n_2\}}^{\min\{n_1,m\}}\binom{n_1}{k}\binom{n_2}{m-k} = \binom{n_1+n_2}{m}. \tag{6.4}$$

Πράγματι, έστω το εξής πρόβλημα: Έχουμε $n_1 > 0$ άντρες και $n_2 > 0$ γυναίκες, και πρέπει να φτιάξουμε μια επιτροπή από m άτομα, με $0 \leq m \leq n_1 + n_2$. Το δεξί μέλος της (6.4) μας δίνει τους δυνατούς συνδυασμούς. Επιπλέον, κάθε ένας από τους όρους του αριστερού σκέλους της (6.4) μας δίνει τους τρόπους με τους οποίους μπορούμε να φτιάξουμε μια επιτροπή m ατόμων με k άντρες. Σχετικά με τα όρια του αθροίσματος, παρατηρήστε ότι σε κάθε όρο του αθροίσματος ο αριθμός των αντρών k θα πρέπει να είναι:

1. το πολύ ίσος με τον ολικό αριθμό αντρών, δηλαδή $k \leq n_1$,

2. το πολύ ίσος με το μέγεθος της επιτροπής, δηλαδή $k \leq m$,

3. μεγαλύτερος ή ίσος του μηδενός, δηλαδή $k \geq 0$, και τέλος,

4. αν $m > n_2$ (δηλαδή τα μέλη της επιτροπής είναι περισσότερα από τις γυναίκες) μεγαλύτερος ή ίσος από το $m - n_2$, δηλαδή $k \geq m - n_2$.

Συνδυάζοντας όλους αυτούς τους περιορισμούς, προκύπτουν τελικά τα όρια του αθροίσματος. Η ταυτότητα (6.4) είναι μια από τις απλούστερες μορφές της ταυτότητα του Vandermonde (και όχι Voldemort).

Εφαρμόζοντας τελικά την (6.4) στην (6.3), προκύπτει τελικά πως

$$p_{X+Y}(m) = \binom{n_1 + n_2}{m} p^m (1-p)^{n_1+n_2-m},$$

δηλαδή το ζητούμενο.

Το αποτέλεσμα είναι αναμενόμενο. Πράγματι, το X εκφράζει το πλήθος των επιτυχιών σε n_1 πειράματα, ανεξάρτητα μεταξύ τους και με πιθανότητα επιτυχίας p, και το Y παρομοίως εκφράζει το πλήθος των επιτυχιών σε n_2 πειράματα, ανεξάρτητα μεταξύ τους, και ανεξάρτητα από τα πρώτα, επίσης με πιθανότητα επιτυχίας p. Άρα, λογικά το $X+Y$ εκφράζει τον πλήθος των επιτυχιών στο σύνολο των n_1+n_2 όμοιων πειραμάτων, και συνεπώς πρέπει να ακολουθεί την διωνυμική κατανομή.

Παράδειγμα 6.25. *(Dungeons & Dragons)* Θα υπολογίσουμε την μάζα του αθροίσματος των αποτελεσμάτων τριών δίκαιων και ανεξάρτητων ζαριών.

Έστω πως X, Y, Z τα αποτελέσματα των τριών ρίψεων. Αρχικά, παρατηρούμε πως

με χρήση του Παραδείγματος 6.22 έχουμε

$$\begin{aligned}
p_{X+Y}(2) &= p_X(1)p_Y(1) = \frac{1}{6} \times \frac{1}{6}, \\
p_{X+Y}(3) &= p_X(2)p_Y(1) + p_X(1)p_Y(2) = 2 \times \frac{1}{6} \times \frac{1}{6}, \\
p_{X+Y}(4) &= p_X(3)p_Y(1) + p_X(2)p_Y(2) + p_X(1)p_Y(3) = 3 \times \frac{1}{6} \times \frac{1}{6}, \\
p_{X+Y}(5) &= p_X(4)p_Y(1) + p_X(3)p_Y(2) + p_X(2)p_Y(3) + p_X(1)p_Y(4) = 4 \times \frac{1}{6} \times \frac{1}{6},
\end{aligned}$$

και συνεχίζοντας με αυτόν τον τρόπο προκύπτουν οι εξής τιμές της μάζας του $X+Y$:

$$\begin{gathered}
p_{X+Y}(2) = \frac{1}{36}, \quad p_{X+Y}(3) = \frac{2}{36}, \quad p_{X+Y}(4) = \frac{3}{36}, \quad p_{X+Y}(5) = \frac{4}{36}, \\
p_{X+Y}(6) = \frac{5}{36}, \quad p_{X+Y}(7) = \frac{6}{36}, \quad p_{X+Y}(8) = \frac{5}{36}, \quad p_{X+Y}(9) = \frac{4}{36}, \\
p_{X+Y}(10) = \frac{3}{36}, \quad p_{X+Y}(11) = \frac{2}{36}, \quad p_{X+Y}(12) = \frac{1}{36}.
\end{gathered}$$

Στα άνω, ουσιαστικά απαριθμήσαμε όλους τους τρόπους με τους οποίους μπορεί σε κάθε περίπτωση να προκύψει ένα αποτέλεσμα για το άθροισμα και υπολογίσαμε το άθροισμα των αντίστοιχων πιθανοτήτων, χρησιμοποιώντας και την ανεξαρτησία των ρίψεων.

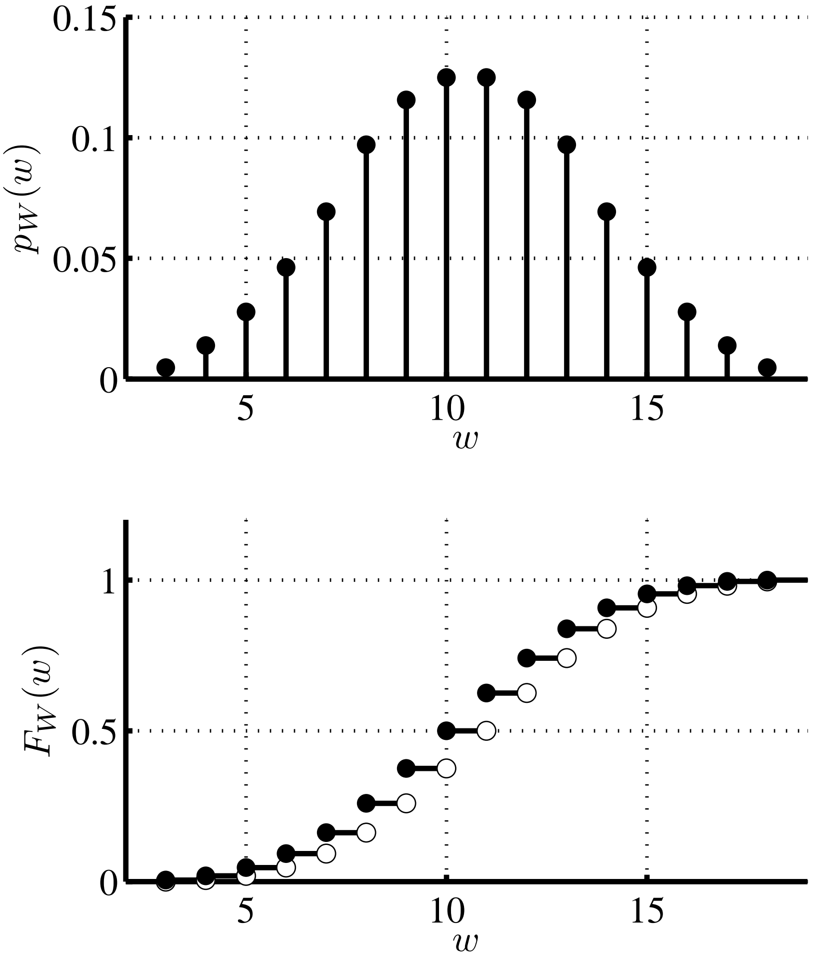
Κατόπιν, για να υπολογίσουμε την μάζα της $X+Y+Z$, επαναλαμβάνουμε την άνω μέθοδο, αλλά για τις τυχαίες μεταβλητές $(X+Y)$ και Z. Για παράδειγμα,

$$\begin{aligned}
p_{X+Y+Z}(3) &= p_{X+Y}(2)p_Z(1) = \frac{1}{36} \times \frac{1}{6} = \frac{1}{216}, \\
p_{X+Y+Z}(4) &= p_{X+Y}(3)p_Z(1) + p_{X+Y}(2)p_Z(2) = \frac{2}{36} \times \frac{1}{6} + \frac{1}{36} \times \frac{1}{6} = \frac{3}{216},
\end{aligned}$$

και συνεχίζοντας με τον ίδιο τρόπο προκύπτει τελικά

$$\begin{gathered}
p_{X+Y+Z}(3) = \frac{1}{216}, \quad p_{X+Y+Z}(4) = \frac{3}{216}, \quad p_{X+Y+Z}(5) = \frac{6}{216}, \\
p_{X+Y+Z}(6) = \frac{10}{216}, \quad p_{X+Y+Z}(7) = \frac{15}{216}, \quad p_{X+Y+Z}(8) = \frac{21}{216}, \\
p_{X+Y+Z}(9) = \frac{25}{216}, \quad p_{X+Y+Z}(10) = \frac{27}{216}, \quad p_{X+Y+Z}(11) = \frac{27}{216}, \\
p_{X+Y+Z}(12) = \frac{25}{216}, \quad p_{X+Y+Z}(13) = \frac{21}{216}, \quad p_{X+Y+Z}(14) = \frac{15}{216}, \\
p_{X+Y+Z}(15) = \frac{10}{216}, \quad p_{X+Y+Z}(16) = \frac{6}{216}, \quad p_{X+Y+Z}(17) = \frac{3}{216}, \\
p_{X+Y+Z}(18) = \frac{1}{216}.
\end{gathered}$$

Η μάζα και η συνάρτηση κατανομής του $X+Y+Z$ έχουν σχεδιαστεί στο Σχήμα 6.1.



Σχήμα 6.1: Η μάζα και η κατανομή του Παραδείγματος 6.25.

6.6 Πολλές Τυχαίες Μεταβλητές

Στις προηγούμενες παραγράφους αυτού του κεφαλαίου περιοριστήκαμε στην περίπτωση δύο από κοινού κατανεμημένων διακριτών Τ.Μ. Όμως όλη η θεωρία που αναπτύξαμε μπορεί να επεκταθεί εύκολα στην περίπτωση πολλών διακριτών Τ.Μ. Σε αυτή την παράγραφο παρουσιάζουμε συνοπτικά την επέκταση της θεωρίας, και τα πιο χρήσιμα αποτελέσματα της. Παραλείπουμε τις περισσότερες αποδείξεις, καθώς είναι όμοιες με αυτές της περίπτωσης των δύο Τ.Μ.

Ορισμός 6.5. (Από κοινού μάζα και περιθώρια μάζα) *Έστω n διακριτές τυχαίες μεταβλητές $X_1, \ldots, X_n$, με σύνολα τιμών $S_{X_1}, \ldots, S_{X_n}$ αντιστοίχως.*

1. *Ορίζουμε την <u>από κοινού συνάρτηση μάζας πιθανότητας</u> ή <u>από κοινού μάζα</u> ως την συνάρτηση $p_{X_1 \ldots X_n}(x_1, \ldots, x_n) : S_{X_1} \times \cdots \times S_{X_n} \to [0,1]$ με*
$$p_{X_1 \ldots X_n}(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n), \quad \forall x_1 \in S_{X_1}, \ldots, \ \forall x_n \in S_{X_n}.$$
2. *Οι μάζες πιθανότητας των $X_1, \ldots, X_n$ καλούνται <u>περιθώριες μάζες πιθανότητας</u>.*

Λήμμα 6.6. (Ιδιότητες από κοινού μάζας) *Έστω n διακριτές Τ.Μ. $X_1, \ldots, X_n$, με σύνολα τιμών $S_{X_1}, \ldots, S_{X_n}$ και μάζες $p_{X_1}(x_1), \ldots, p_{X_n}(x_n)$ αντιστοίχως. Η από κοινού μάζα τους $p_{X_1 \ldots X_n}(x_1, \ldots, x_n)$ έχει τις εξής ιδιότητες:*

1. $\displaystyle\sum_{x_1 \in S_{X_1}, \ldots, x_n \in S_{X_n}} p_{X_1 \ldots X_n}(x_1, \ldots, x_n) = 1.$
2. *Έστω A ένα οποιοδήποτε υποσύνολο του καρτεσιανού γινομένου $S_{X_1} \times \cdots \times S_{X_n}$. Τότε*
$$P((X_1, \ldots, X_n) \in A) = \sum_{(x_1, \ldots, x_n) \in A} p_{X_1 \ldots X_n}(x_1, \ldots, x_n). \tag{6.5}$$
3. $\displaystyle p_{X_i}(x) = \sum_{x_1 \in S_{X_1}, \ldots, x_{i-1} \in S_{X_{i-1}}, x_{i+1} \in S_{X_{i+1}}, \ldots, x_n \in S_{X_n}} p_{X_1 \ldots X_n}(x_1, \ldots, x_n).$

Λήμμα 6.7. (Ιδιότητες μέσης τιμής συνάρτησης πολλών Τ.Μ.) *Έστω n διακριτές τυχαίες μεταβλητές $X_1, \ldots, X_n$, με σύνολα τιμών $S_{X_1}, \ldots, S_{X_n}$ και από κοινού μάζα $p_{X_1 \ldots X_n}(x_1, \ldots, x_n)$.*

1. *Έστω Τ.Μ. $Z = g(X_1, \ldots, X_n)$. Η μέση τιμή της ισούται με*
$$E(Z) = E(g(X_1, \ldots, X_n)) = \sum_{x_1 \in S_{X_1}, \ldots, x_n \in S_{X_n}} g(x_1, \ldots, x_n) p_{X_1 \ldots X_n}(x_1, \ldots, x_n).$$

2. *Αν* $a_i, b \in \mathbb{R}$,

$$E\left(\sum_{i=1}^{n} a_i X_i + b\right) = \sum_{i=1}^{n} a_i E(X_i) + b.$$

3. *Πιο γενικά, αν έχουμε* K *Τ.Μ.* $Z_k = g_k(X_1, \ldots, X_n)$, $k = 1, \ldots, K$, *και* $a_k \in \mathbb{R}$,

$$E\left(\sum_{k=1}^{K} a_k g_k(X_1, \ldots, X_n)\right) = \sum_{k=1}^{K} a_k E(g_k(X_1, \ldots, X_n)).$$

4.

$$\mathrm{VAR}\left(\sum_{i=1}^{n} X_i + b\right) = \sum_{i=1}^{n} \mathrm{VAR}(X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{COV}(X_i, X_j). \tag{6.6}$$

Απόδειξη. Θα αποδείξουμε μόνο το τελευταίο σκέλος, καθώς παρουσιάζει ενδιαφέρον. Παρατηρήστε πως

$$\begin{aligned}
&\mathrm{VAR}\left(\sum_{i=1}^{n} X_i + b\right) \\
&= \mathrm{VAR}\left(\sum_{i=1}^{n} X_i\right) = E\left(\left(\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)\right)^2\right) \\
&= E\left(\left(\sum_{i=1}^{n} X_i - E\left(\sum_{i=1}^{n} X_i\right)\right) \times \left(\sum_{j=1}^{n} X_j - E\left(\sum_{j=1}^{n} X_j\right)\right)\right) \\
&= E\left(\left(\sum_{i=1}^{n} (X_i - E(X_i))\right) \times \left(\sum_{j=1}^{n} (X_j - E(X_j))\right)\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} E\left((X_i - E(X_i))(X_j - E(X_j))\right) \\
&= \sum_{i=1}^{n} E\left((X_i - E(X_i))(X_i - E(X_i))\right) \\
&\quad + \sum_{i=1}^{n} \sum_{j=1,\ j \ne i}^{N} E\left((X_i - E(X_i))(X_j - E(X_j))\right) \\
&= \sum_{i=1}^{n} \mathrm{VAR}(X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{COV}(X_i, X_j).
\end{aligned}$$

□

Ορισμός 6.6. (n ανεξάρτητες Τ.Μ.) *n διακριτές Τ.Μ. $X_1, \ldots, X_n$, με σύνολα τιμών $S_{X_1}, \ldots, S_{X_n}$, αντιστοίχως, καλούνται* <u>*ανεξάρτητες*</u> *αν για οποιαδήποτε υποσύνολα $A_i \subseteq S_{X_i}$, $i = 1, \ldots, n$ ισχύει*

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \ldots P(X_n \in A_n).$$

Λήμμα 6.8. (Κριτήριο ανεξαρτησίας Τ.Μ.) *Έστω n διακριτές Τ.Μ. $X_1, \ldots, X_n$, με σύνολα τιμών $S_{X_1}, \ldots, S_{X_n}$, από κοινού μάζα $p_{X_1 \ldots X_n}(x_1, \ldots, x_n)$, και περιθώριες μάζες $p_{X_i}(x_i)$, $i = 1, \ldots, n$. Οι $X_1, \ldots, X_n$ είναι ανεξάρτητες αν και μόνο αν για οποιαδήποτε $x_i \in S_{X_i}$ έχουμε*

$$p_{X_1 \ldots X_n}(x_1, \ldots, x_n) = p_{X_1}(x_1) \ldots p_{X_n}(x_n).$$

Λήμμα 6.9. (Ιδιότητες ανεξάρτητων μεταβλητών) *Έστω n διακριτές ανεξάρτητες Τ.Μ. $X_1, \ldots, X_n$.*

1. *Οποιοδήποτε υποσύνολο από τις $X_1, \ldots, X_n$ είναι επίσης ανεξάρτητες Τ.Μ.*
2. *Έστω συναρτήσεις $g_i : S_{X_i} \to \mathbb{R}$. Θα ισχύει*

$$E(g_1(X_1) \ldots g_n(X_n)) = E(g_1(X_1)) \ldots E(g_n(X_n)).$$

Ειδική περίπτωση της άνω είναι η

$$E(X_1 \ldots X_n) = E(X_1) \ldots E(X_n).$$

3. $\mathrm{VAR}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{VAR}(X_i).$

Παράδειγμα 6.26. Έστω τρεις ανεξάρτητες Τ.Μ. X_1, X_2, X_3 που έχουν όλες την ίδια μάζα, και, κατά συνέπεια, την ίδια μέση τιμή $\mu = E(X_i)$. Ορίζουμε δύο νέες Τ.Μ., τις $Y = X_1 + X_2$ και $Z = X_1 + X_3$. Εφόσον και οι δύο εξαρτώνται απ' την X_1 περιμένουμε πως δεν θα είναι ανεξάρτητες, και πως πιθανότατα η συνδιακύμανσή τους δεν θα είναι μηδενική. Έχουμε:

$$\begin{aligned}
\mathrm{COV}(Y, Z) &= E(YZ) - E(Y)E(Z) \\
&= E((X_1 + X_2)(X_1 + X_3)) - E(X_1 + X_2)E(X_1 + X_3) \\
&= E(X_1^2 + X_1X_3 + X_1X_2 + X_2X_3) - E(X_1 + X_2)E(X_1 + X_3) \\
&= E(X_1^2) + E(X_1X_3) + E(X_1X_2) + E(X_2X_3) \\
&\qquad -(E(X_1))^2 - E(X_1)E(X_3) - E(X_1)E(X_2) - E(X_2)E(X_3).
\end{aligned}$$

Όμως επιπλέον, λόγω της ανεξαρτησίας των X_i, έχουμε $E(X_iX_j) = E(X_i)E(X_j)$. Συνεπώς, οι έξι όροι πιο πάνω εκτός του πρώτου και του τέταρτου απλοποιούνται, διότι είναι όλοι ίσοι με $\pm(E(X_i))^2 = \mu^2$, και τελικά προκύπτει πως $\mathrm{COV}(Y, Z) = \mathrm{VAR}(X_1)$. Άρα, η συνδιακύμανση μεταξύ Y και Z όχι μόνο είναι μη μηδενική (όπως περιμέναμε), αλλά επιπλέον ισούται με τη διασπορά της Τ.Μ. X_1 η οποία συνδέει τις Y και Z.

Παράδειγμα 6.27. *(Μέση τιμή και διασπορά της διωνυμικής κατανομής)* Θα χρησιμοποιήσουμε την θεωρία που έχουμε αναπτύξει μέχρι τώρα για να υπολογίσουμε, με πολύ εύκολο τρόπο, την μέση τιμή και τη διασπορά της διωνυμικής κατανομής. Έστω Τ.Μ. X που ακολουθεί την διωνυμική κατανομή με παραμέτρους N και p. Κατά τα γνωστά από τη θεωρία, η X εκφράζει τον αριθμό των επιτυχιών σε N ανεξάρτητα πειράματα, καθένα με πιθανότητα επιτυχίας p. Συνεπώς, η X μπορεί να γραφτεί ως $X = \sum_{i=1}^{N} X_i$ όπου οι X_i, $i = 1, \ldots, N$ είναι ανεξάρτητες Τ.Μ. Bernoulli, όλες με παράμετρο p, και συνεπώς με μέση τιμή $E(X_i) = p$ και διασπορά $\mathrm{VAR}(X_i) = p(1-p)$. Έχουμε:

$$E(X) = E\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} E(X_i) = \sum_{i=1}^{N} p = Np.$$

Για να υπολογίσουμε τη διασπορά, παρατηρούμε απλώς πως

$$\mathrm{VAR}(X) = \mathrm{VAR}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} \mathrm{VAR}(X_i) = Np(1-p).$$

Παράδειγμα 6.28. *(Μέση τιμή και διασπορά της υπεργεωμετρικής κατανομής)* Με κάποιες τροποποιήσεις, μπορούμε να χρησιμοποιήσουμε τη μέθοδο του προηγούμενου παραδείγματος για να υπολογίσουμε τη μέση τιμή και τη διασπορά της υπεργεωμετρικής κατανομής. Έστω X με κατανομή $\text{Υπερ}(N, k, n)$. Η X εκφράζει τον αριθμό αντικειμένων τύπου I που θα λάβουμε αν πάρουμε n αντικείμενα από N, εκ των οποίων k είναι τύπου I και τα υπόλοιπα τύπου II. Συνεπώς, η X μπορεί να εκφραστεί ως άθροισμα $X = \sum_{i=1}^{n} X_i$, όπου οι X_i είναι όλες Bernoulli, προφανώς με παράμετρο $\frac{k}{N}$. Άρα:

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{k}{N} = \frac{nk}{N}.$$

Για να υπολογίσουμε τη διασπορά, παρατηρούμε καταρχήν πως οι X_i ΔΕΝ είναι ανεξάρτητες. Για παράδειγμα, αν το πρώτο αντικείμενο που θα πάρουμε είναι τύπου I, τότε μειώνεται κάπως η πιθανότητα να είναι τύπου I και το δεύτερο αντικείμενο. Άρα, θα πρέπει να χρησιμοποιήσουμε τη σχέση

$$\mathrm{VAR}(X) = \mathrm{VAR}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{VAR}(X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{COV}(X_i, X_j).$$

Παρατηρούμε καταρχήν ότι, αφού οι X_i είναι όλες Bernoulli με παράμετρο $\frac{k}{N}$, θα έχουμε

$$\mathrm{VAR}(X_i) = \frac{k}{N}\left(1 - \frac{k}{N}\right).$$

Επιπλέον, προφανώς, για λόγους συμμετρίας, θα έχουμε επιπλέον

$$\mathrm{COV}(X_i, X_j) = \mathrm{COV}(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2),$$

για κάθε i, j τέτοια ώστε $1 \leq i < j \leq n$. Παρατηρούμε όμως πως η Τ.Μ. X_1X_2 είναι 1 αν και μόνο αν τα πρώτα δύο αντικείμενα είναι τύπου Ι. Αυτό γίνεται με πιθανότητα $\frac{k}{N} \times \frac{k-1}{N-1}$. Άρα τελικά

$$E(X_1X_2) = 1 \times \frac{k}{N} \times \frac{k-1}{N-1} + 0 \times \left(1 - \frac{k}{N} \times \frac{k-1}{N-1}\right) = \frac{k(k-1)}{N(N-1)},$$

και συνδυάζοντας όλα τα άνω:

$$\begin{aligned}\mathrm{VAR}(X) &= n\frac{k}{N}\left(1 - \frac{k}{N}\right) + n(n-1)\left[\frac{k(k-1)}{N(N-1)} - \frac{k^2}{N^2}\right] \\ &= \frac{nk}{N}\left[1 - \frac{k}{N} + \frac{(n-1)(k-1)}{N-1} - \frac{(n-1)k}{N}\right] = \frac{nk(N-k)(N-n)}{N^2(N-1)}.\end{aligned}$$

Παράδειγμα 6.29. *(Το πρόβλημα των συζύγων)* Σε κάποιο πάρτι, N άτομα αφήνουν τα καπέλα τους σε ένα τραπέζι. Τα καπέλα ανακατεύονται και, φεύγοντας, κάθε άτομο παίρνει στην τύχη ένα καπέλο. Έστω X το πλήθος των ατόμων που παίρνουν το δικό τους καπέλο. Θα υπολογίσουμε τη μέση τιμή και τη διασπορά της X.

Θα χρησιμοποιήσουμε τις βοηθητικές Τ.Μ. X_i, $i = 1, 2, \ldots, N$, όπου $X_i = 1$ αν το άτομο i πάρει το δικό του καπέλο, και $X_i = 0$ στην αντίθετη περίπτωση. Παρατηρούμε ότι $X = X_1 + X_2 + \cdots + X_N$. Επιπλέον, λόγω της συμμετρίας του προβλήματος, η πιθανότητα να πάρει ένα άτομο το δικό του καπέλο είναι $\frac{1}{N}$. Συνεπώς η μέση τιμή της κάθε X_i είναι

$$E(X_i) = 1 \times \frac{1}{N} + 0 \times (1 - \frac{1}{N}) = \frac{1}{N},$$

και συνεπώς

$$E(X) = E(X_1 + X_2 + \cdots + X_N) = E(X_1) + E(X_2) + \cdots + E(X_N) = N \times \frac{1}{N} = 1.$$

Άρα, κατά μέσο όρο, ακριβώς ένα άτομο θα φύγει με το δικό του καπέλο. Παρατηρήστε ότι το αποτέλεσμα είναι ανεξάρτητο του N, κάτι που ίσως να μην αναμενόταν.

Για να υπολογίσουμε τη διασπορά, μπορούμε να χρησιμοποιήσουμε την (6.6). Εναλλακτικά (και με περίπου τον ίδιο όγκο πράξεων) παρατηρούμε καταρχήν πως

$$\begin{aligned} E(X^2) &= E\left(\left(\sum_{i=1}^{N} X_i\right)^2\right) = E\left(\left(\sum_{i=1}^{N} X_i\right)\left(\sum_{j=1}^{N} X_j\right)\right) \\ &= E\left(\sum_{i=1}^{N}\sum_{j=1}^{N} X_i X_j\right) = \sum_{i=1}^{N}\sum_{j=1}^{N} E(X_i X_j). \end{aligned}$$

Στο τελευταίο πιο πάνω διπλό άθροισμα, υπάρχουν N όροι της μορφής $E(X_i^2)$, και $N(N-1)$ όροι της μορφής $E(X_i X_j)$ για $i \neq j$. Εφόσον οι X_i είναι Τ.Μ. Bernoulli, πάντοτε έχουμε $X_i^2 = X_i$ και συνεπώς,

$$E(X_i^2) = E(X_i) = \frac{1}{N},$$

ενώ για $i \neq j$, βρίσκουμε, από τον ορισμό της δεσμευμένης πιθανότητας,

$$\begin{aligned} E(X_i X_j) &= P(X_i = X_j = 1) \times 1 + (1 - P(X_i = X_j = 1)) \times 0 \\ &= P(X_i = 1 \text{ και } X_j = 1) \\ &= P(X_i = 1) P(X_j = 1 | X_i = 1) = \frac{1}{N} \times \frac{1}{N-1}. \end{aligned}$$

Άρα, τελικά,

$$E(X^2) = E\left(\left(\sum_{i=1}^{N} X_i\right)^2\right) = N\frac{1}{N} + N(N-1)\frac{1}{N(N-1)} = 2,$$

και για τη διασπορά έχουμε

$$\mathrm{VAR}(X) = E(X^2) - (E(X))^2 = 1.$$

Αν και δεν το αποδεικνύουμε εδώ, αναφέρουμε ότι η κατανομή του X, καθώς το $N \to \infty$, τείνει στην κατανομή Poisson με παράμετρο $\lambda = 1$. Μπορείτε να εξηγήσετε διαισθητικά γιατί;

Παράδειγμα 6.30. *(Επιζώντα ζευγάρια)* Έστω μία πόλη με N ζεύγη ανδρών και γυναικών (δηλαδή με συνολικό πληθυσμό $2N$ ατόμων.) Αν υποθέσουμε ότι k τυχαία άτομα πεθαίνουν, θα υπολογίσουμε τη μέση τιμή και η διασπορά του πλήθους των ζευγαριών που παραμένουν. Εννοείται ότι όλοι οι άνθρωποι είναι εξίσου δυνατόν να πεθάνουν, και οι θάνατοι είναι ανεξάρτητοι μεταξύ τους. Το πρόβλημα αυτό διατυπώθηκε για πρώτη φορά από τον Daniel Bernoulli, το 1768.

Έστω οι τυχαίες μεταβλητές M_i, F_i, $i = 1, 2, \ldots, N$, όπου, για κάθε i, έχουμε $M_i = 1$ (ή 0) αν ο i-οστός άνδρας επιβιώσει (ή πεθάνει), και αντιστοίχως $F_i = 1$ (ή 0) αν η i-οστή γυναίκα επιβιώσει (ή πεθάνει). Επιπλέον, ορίζουμε τις Τ.Μ. Z_i οι οποίες παίρνουν την τιμή 1 αν το ζευγάρι i παραμείνει, αλλιώς $Z_i = 0$, και παρατηρούμε πως $Z_i = M_i F_i$, για κάθε i. Το συνολικό πλήθος, έστω Z, των ζευγαριών που παραμένουν είναι $Z = \sum_{i=1}^N Z_i = \sum_{i=1}^N M_i F_i$ και για τον υπολογισμό της ζητούμενης μέσης τιμής $E(Z)$ αρκεί να υπολογίσουμε την $E(Z_i)$ της κάθε Z_i. Παρατηρούμε ότι, από τον ορισμό της μέσης τιμής και τον ορισμό της δεσμευμένης πιθανότητας, έχουμε,

$$\begin{aligned} E(Z_i) &= 1 \times P(Z_i = 1) + 0 \times P(Z_i = 0) = P(\{M_i = 1\} \cap \{F_i = 1\}) \\ &= P(M_i = 1) P(F_i = 1 | M_i = 1) = \frac{2N - k}{2N} \times \frac{2N - k - 1}{2N - 1}, \end{aligned}$$

όπου η τελευταία ισότητα προκύπτει καθώς, αν γνωρίζουμε ότι επιβίωσε ο άνδρας i, μένουν $2N - 1$ άτομα εκ των οποίων θα επιβιώσουν τα $2N - k - 1$. Συνδυάζοντας τα πιο πάνω βρίσκουμε την ζητούμενη μέση τιμή:

$$E(Z) = E\left(\sum_{i=1}^N Z_i\right) = \sum_{i=1}^N E(Z_i) = \frac{(2N - k)(2N - k - 1)}{2(2N - 1)}.$$

Επιπλέον,

$$\begin{aligned} E(Z_i^2) &= 1^2 \times P(M_i = F_i = 1) + 0^2 \times P((M_i = F_i = 1)') = P(M_i = F_i = 1) \\ &= \frac{2N - k}{2N} \times \frac{2N - k - 1}{2N - 1}. \end{aligned}$$

Ακολούθως, παρατηρούμε ότι

$$E(Z^2) = E\left(\left(\sum_{i=1}^N Z_i\right)^2\right) = E\left(\left(\sum_{i=1}^N Z_i\right)\left(\sum_{j=1}^N Z_j\right)\right) = \sum_{i=1}^N \sum_{j=1}^N E(Z_i Z_j).$$

Όμως, αν $i \neq j$:

$$\begin{aligned} E(Z_i Z_j) &= P(M_i = M_j = F_i = F_j = 1) \\ &= \frac{\binom{2N-4}{k}}{\binom{2N}{k}} = \frac{(2N - k)(2N - k - 1)(2N - k - 2)(2N - k - 3)}{N(N - 1)(N - 2)(N - 3)}. \end{aligned}$$

Η τελευταία ισότητα προκύπτει καθώς υπάρχουν $\binom{2N}{k}$ συνδυασμοί θανάτων, αλλά μόνο $\binom{2N-4}{k}$ από αυτούς αφήνουν άθικτα τα δύο ζεύγη.

Συνδυάζοντας τις άνω, προκύπτει, μετά από πράξεις, ότι:

$$\begin{aligned} \mathrm{VAR}(Z) = E(Z^2) - (E(Z))^2 &= (2N - k)(2N - k - 1) \\ &\times \left[\frac{(2N - k - 2)(2N - k - 3)}{(N - 2)(N - 3)} + \frac{1}{2(2N - 1)} - \frac{(2N - k)(2N - k - 1)}{4(2N - 1)^2}\right]. \end{aligned}$$

Κεφάλαιο 7

Συνεχείς Τυχαίες Μεταβλητές

Στα προηγούμενα κεφάλαια ασχοληθήκαμε αποκλειστικά με τη μια από τις δύο σημαντικές κατηγορίες τυχαίων μεταβλητών, τις διακριτές. Σε αυτό το κεφάλαιο θα ασχοληθούμε με τη δεύτερη σημαντική κατηγορία, τις συνεχείς.

Δύο βασικοί λόγοι ωθούν το ενδιαφέρον μας για τις συνεχείς Τ.Μ. Ο ένας είναι προφανής: Πολλές ποσότητες που είναι σημαντικές στην πράξη, μπορούν από τη φύση τους να μοντελοποιηθούν εύκολα και επακριβώς ως συνεχείς Τ.Μ. — π.χ., ο χρόνος που διαρκεί η εκτέλεση ενός αλγορίθμου, η θερμοκρασία ενός επεξεργαστή, η απόσταση μεταξύ ενός κινητού τηλεφώνου και της κεραίας με την οποία επικοινωνεί, κ.ο.κ. Ο δεύτερος λόγος είναι πιο λεπτός, και σχετίζεται με Κεντρικό Οριακό Θεώρημα. Σύμφωνα με το αυτό το θεώρημα, αν οι Τ.Μ. $X_1, X_2, \ldots, X_N$ είναι ανεξάρτητες και έχουν όλες την ίδια κατανομή τότε ο *εμπειρικός μέσος όρος*

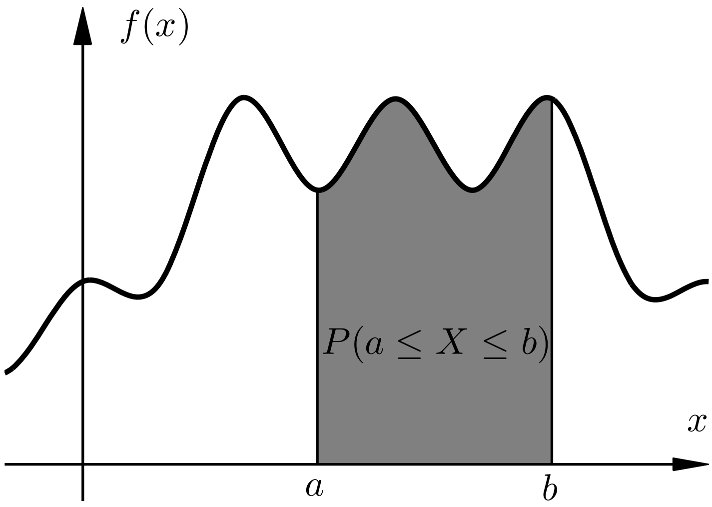
$$\bar{X}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$$

μπορεί να προσεγγισθεί μέσω της κατανομής μιας *συνεχούς* τυχαίας μεταβλητής, συγκεκριμένα της λεγόμενης *κανονικής κατανομής*. Αυτό ισχύει ακόμα και αν οι Τ.Μ. $X_1, X_2, \ldots, X_N$ δεν είναι συνεχείς.

7.1 Συνεχείς Τυχαίες Μεταβλητές

Ορισμός 7.1. (Συνεχείς Τ.Μ.) *Συνεχής Τυχαία Μεταβλητή (Τ.Μ.) καλείται κάθε συνάρτηση $X : \Omega \to \mathbb{R}$ όπου ο Ω είναι ένας δειγματικός χώρος, και για την οποία υπάρχει μια συνάρτηση $f : \mathbb{R} \to [0, \infty)$, η οποία καλείται συνάρτηση πυκνότητας πιθανότητας, ή πυκνότητα πιθανότητας, ή απλώς πυκνότητα, τέτοια ώστε για κάθε $B \subseteq \mathbb{R}$*

$$P(X \in B) = \int_B f(x)\, dx. \tag{7.1}$$



Σχήμα 7.1: Γραφική ερμηνεία του υπολογισμού της πιθανότητας $P(a \leq X \leq b)$ για μια συνεχή Τ.Μ. X μέσω της πυκνότητας $f(x)$. Η πιθανότητα $P(a \leq X \leq b)$ είναι ίση με το εμβαδόν του σκιασμένου χωρίου που σχηματίζεται από το γράφημα της $f(x)$, τον άξονα x, και τις ευθείες $x = a$ και $x = b$.

Παρατήρηση: Στην απλούστερη περίπτωση χρήσης της (7.1), το B είναι κλειστό και φραγμένο διάστημα και η $f(x)$ ολοκληρώσιμη σε αυτό. Τότε το ολοκλήρωμα υπάρχει με την αυστηρή έννοια του όρου. Για παράδειγμα, αν $B = [a, b]$, και η $f(x)$ είναι ολοκληρώσιμη στο $[a, b]$, η (7.1) γίνεται

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx.$$

Η ολοκλήρωση απεικονίζεται γραφικά στο Σχήμα 7.1. Δείτε το ακόλουθο παράδειγμα.

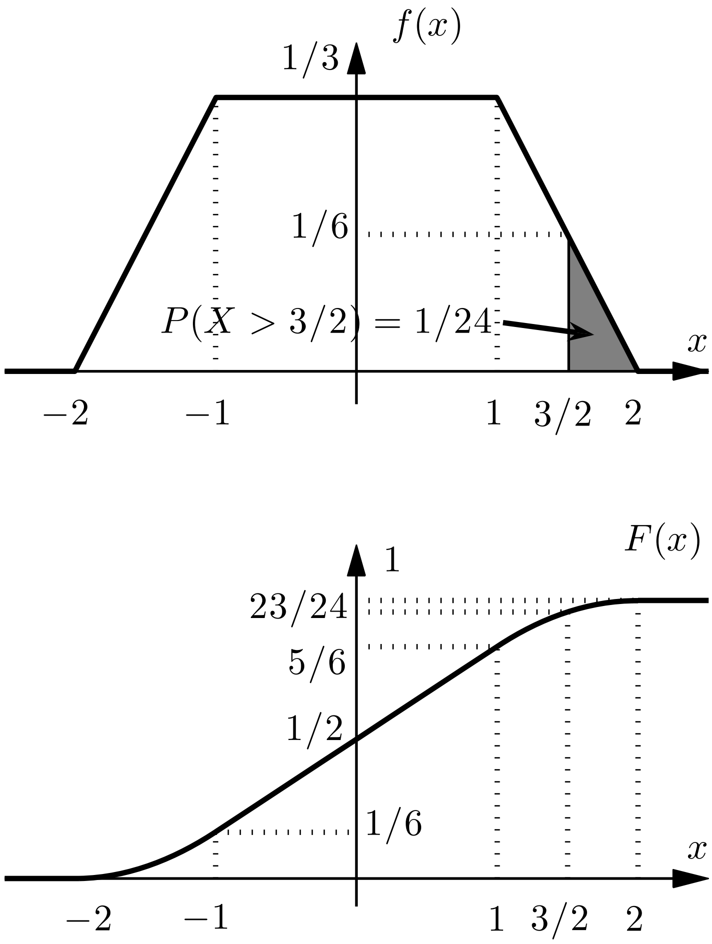
Παράδειγμα 7.1. *(Τμηματικά γραμμική πυκνότητα πιθανότητας)* Μια συνεχής Τ.Μ. X έχει την πυκνότητα

$$f(x) = \begin{cases} 0, & x < -2, \\ \frac{1}{3}(2 + x), & -2 \leq x < -1, \\ \frac{1}{3}, & -1 \leq x < 1, \\ \frac{1}{3}(2 - x), & 1 \leq x < 2, \\ 0, & x \geq 2. \end{cases}$$

Θα απαντήσουμε τα ακόλουθα ερωτήματα:

1. Ποια είναι η πιθανότητα $P\left(\frac{3}{2} \leq X \leq 2\right)$;
2. Ποια είναι η πιθανότητα $P\left(\frac{3}{2} \leq X \leq 2 | 0 \leq X \leq 2\right)$;

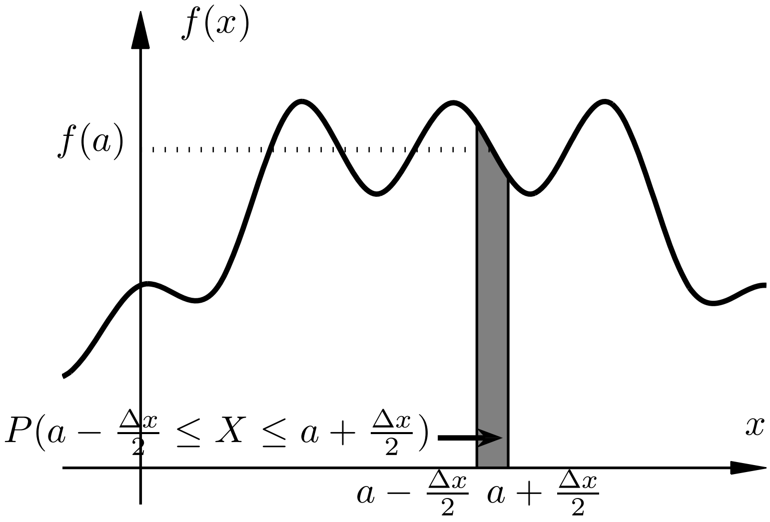
Έχουμε, κατά περίπτωση,



Σχήμα 7.2: Παραδείγματα 7.1, 7.2.

1.
$$P\left(\frac{3}{2} \le X \le 2\right) = \int_{3/2}^{2} \frac{1}{3}(2-x)\,dx = \left[\frac{2x}{3}\right]_{3/2}^{2} - \left[\frac{x^2}{6}\right]_{3/2}^{2} = \frac{1}{3} - \frac{7}{24} = \frac{1}{24}.$$

2. Από τον ορισμό της δεσμευμένης πιθανότητας, έχουμε
$$P\left(\frac{3}{2} \le X \le 2 | 0 \le X \le 2\right) = \frac{P\left(\frac{3}{2} \le X \le 2,\ 0 \le X \le 2\right)}{P(0 \le X \le 2)} = \frac{P\left(\frac{3}{2} \le X \le 2\right)}{P(0 \le X \le 2)},$$
όπου $P(\frac{3}{2} \le X \le 2) = \frac{1}{24}$ και
$$P(0 \le X \le 2) = \int_0^1 \frac{1}{3}\,dx + \int_1^2 \frac{1}{3}(2-x)\,dx = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}.$$
Άρα
$$P\left(\frac{3}{2} \le X \le 2 | 0 \le X \le 2\right) = \left[\frac{1}{24}\right] \times \left[\frac{1}{2}\right]^{-1} = \frac{1}{12}.$$



Σχήμα 7.3: Διαισθητική ερμηνεία της πυκνότητας $f(x)$: η πιθανότητα $P(a - \frac{\Delta x}{2} \leq X \leq a + \frac{\Delta x}{2})$ ισούται με το εμβαδόν του σκιασμένου χωρίου, το οποίο, όταν το Δx είναι μικρό, προσεγγιστικά είναι ίσο με το εμβαδόν $f(a)\Delta x$ ενός ορθογωνίου ύψους $f(a)$ και πλάτους Δx. (Φανταστείτε τι συμβαίνει όταν $\Delta x \to 0$.)

Παρατηρήσεις

1. Διαισθητικά, η πυκνότητα $f(a)$ στη θέση a εκφράζει πόσο μεγάλη είναι η πιθανότητα να λάβει η X τιμές κοντά στην a. Πράγματι, έστω πως η $f(x)$ είναι συνεχής στο a. Τότε, από το Θεμελιώδες Θεώρημα του Ολοκληρωτικού Λογισμού, προκύπτει πως ισχύει το εξής όριο:

$$\lim_{\Delta x \to 0^+} \frac{P\left(a - \frac{\Delta x}{2} \leq X \leq a + \frac{\Delta x}{2}\right)}{\Delta x} = \lim_{\Delta x \to 0^+} \frac{1}{\Delta x} \int_{a-\frac{\Delta x}{2}}^{a+\frac{\Delta x}{2}} f(x)\, dx = f(a),$$

   και συνεπώς

$$\Delta x \simeq 0, \Delta x > 0 \Rightarrow P\left(a - \frac{\Delta x}{2} \leq X \leq a + \frac{\Delta x}{2}\right) \simeq f(a)\Delta x.$$

   Η ιδιότητα αυτή εξηγείται διαισθητικά στο Σχήμα 7.3.

2. Λόγω της φύσης του ολοκληρώματος, μπορεί δύο Τ.Μ. να έχουν ακριβώς την ίδια συμπεριφορά, δηλαδή οι πιθανότητές τους να βρεθούν σε ένα σύνολο να είναι ίδιες για οποιοδήποτε σύνολο, αλλά οι πυκνότητές τους να είναι διαφορετικές! Σκεφτείτε, για παράδειγμα, τι θα γίνει αν αλλάξουμε σε ένα σημείο μια πυκνότητα. Παρομοίως, οι πιθανότητες όλων των ενδεχόμενων που αφορούν μια Τ.Μ. μπορούν να περιγραφούν βάσει περισσότερων από μιας πυκνοτήτων. Στα ακόλουθα, όταν θα λέμε ότι βρήκαμε την πυκνότητα μιας Τ.Μ., θα εννοείται ότι βρήκαμε *μία* πυκνότητα της Τ.Μ.

3. Στον ορισμό της πυκνότητας πιθανότητας δεν περιοριζόμαστε στον αυστηρό ορισμό του ολοκληρώματος:

   (α′) Αν το B δεν είναι κλειστό, τότε το ολοκλήρωμα λαμβάνεται στην κλειστότητα του B, δηλαδή στο μικρότερο κλειστό υπερσύνολο του B. Για παράδειγμα, αν το $B = (1, 2)$, τότε το ολοκλήρωμα λαμβάνεται στο $[1, 2]$.

   (β′) Αν το B είναι ένωση διαστημάτων, τότε το ολοκλήρωμα λαμβάνεται ως το άθροισμα των επί μέρους ολοκληρωμάτων.

   (γ′) Επιπλέον, το ολοκλήρωμα μπορεί να είναι καταχρηστικό. Για παράδειγμα, αν η $B = (-\infty, k)$, η (7.1) γίνεται

$$P(X < k) = \int_{-\infty}^{k} f(x)\,dx.$$

4. Υπενθυμίζουμε τους ακόλουθους ορισμούς καταχρηστικών ολοκληρωμάτων πρώτου τύπου:

$$\begin{aligned}
\int_{-\infty}^{a} f(x)\,dx &\triangleq \lim_{t\to-\infty}\int_{t}^{a} f(x)\,dx,\\
\int_{a}^{\infty} f(x)\,dx &\triangleq \lim_{t\to\infty}\int_{a}^{t} f(x)\,dx,\\
\int_{-\infty}^{\infty} f(x)\,dx &\triangleq \lim_{t\to-\infty}\int_{t}^{c} f(x)\,dx + \lim_{t\to\infty}\int_{c}^{t} f(x)\,dx,
\end{aligned}$$

   εφόσον δεν προκύπτει απροσδιοριστία $\infty - \infty$, και όπου το $c \in \mathbb{R}$ και η επιλογή του δεν έχει σημασία.

5. Υπενθυμίζουμε τους ακόλουθους ορισμούς καταχρηστικών ολοκληρωμάτων δεύτερου τύπου:

$$\begin{aligned}
\int_{a}^{b} f(x)\,dx &\triangleq \lim_{t\to b^-}\int_{a}^{t} f(x)\,dx,\\
\int_{a}^{b} f(x)\,dx &\triangleq \lim_{t\to a^+}\int_{t}^{b} f(x)\,dx,
\end{aligned}$$

   όπου $a < b$. Στην πρώτη περίπτωση η $f(x)$ δεν είναι ολοκληρώσιμη στο $[a, b]$, αλλά είναι στο $[a, t]$ για κάθε $t < b$. (Για παράδειγμα, γιατί $\lim_{t\to b^-} f(x) = \infty$.) Στη δεύτερη περίπτωση η $f(x)$ δεν είναι ολοκληρώσιμη στο $[a, b]$, αλλά είναι στο $[t, b]$ για κάθε $t > a$. (Για παράδειγμα, γιατί $\lim_{t\to a^+} f(x) = \infty$.)

6. Υπενθυμίζουμε πως τα καταχρηστικά ολοκληρώματα μπορούν να συνδυαστούν με διάφορους τρόπους, εφόσον βέβαια δεν προκύπτουν απροσδιοριστίες $\infty - \infty$. Για παράδειγμα, μπορεί ένα ολοκλήρωμα να είναι καταχρηστικό πρώτου τύπου λόγω του αριστερού του άκρου και καταχρηστικό δευτέρου τύπου λόγω του δεξιού του άκρου. Περισσότερα σε εισαγωγικά μαθήματα Ολοκληρωτικού Λογισμού.

7. Ο Ορισμός 7.1 δεν ήταν απολύτως σωστός, γιατί υπάρχουν σύνολα B τόσο πολύπλοκα, που είναι αδύνατο να έχουν ολοκλήρωμα, ακόμα και αν η $f(x)$ είναι πολύ απλή, για παράδειγμα σταθερή. Θα αγνοήσουμε αυτό το πρόβλημα.

8. Υπάρχουν Τ.Μ. που δεν είναι *ούτε* διακριτές, *ούτε* συνεχείς, αλλά σε αυτό το μάθημα μας απασχολούν μόνο αυτά τα δύο είδη.

9. Οποιαδήποτε συνάρτηση $f(x) \geq 0$ έχει ολοκλήρωμα $\int_{-\infty}^{\infty} f(x)\,dx = 1$ μπορεί να θεωρηθεί ότι είναι η πυκνότητα μιας Τ.Μ. (Γιατί χρειάζεται η ισότητα;)

Ορισμός 7.2. *Η συνάρτηση κατανομής πιθανότητας, ή κατανομή πιθανότητας, ή απλώς κατανομή της X είναι η συνάρτηση $F : \mathbb{R} \to [0,1]$ η οποία ορίζεται ως*

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt. \tag{7.2}$$

Παράδειγμα 7.2. Θα υπολογίσουμε την κατανομή της Τ.Μ. X του Παραδείγματος 7.1. Παίρνουμε περιπτώσεις. Αν $x < -2$, τότε

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0.$$

Αν $-2 \leq x \leq -1$,

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-2} 0\,dt + \int_{-2}^{x} \frac{1}{3}(2+t)\,dt = \left[\frac{2t}{3} + \frac{t^2}{6}\right]_{-2}^{x} = \frac{1}{6}(x^2+4x+4).$$

Αν $-1 \leq x \leq 1$, τότε

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{-2} 0\,dt + \int_{-2}^{-1} \frac{1}{3}(2+t)\,dt + \int_{-1}^{x} \frac{1}{3}\,dt \\ &= \left[\frac{2t}{3} + \frac{t^2}{6}\right]_{-2}^{-1} + \left[\frac{t}{3}\right]_{-1}^{x} = \frac{1}{6}(2x+3), \end{aligned}$$

και τελικά, υπολογίζοντας και τις άλλες περιπτώσεις, προκύπτει

$$F(x) = \begin{cases} 0, & x < -2, \\ \frac{1}{6}(x^2 + 4x + 4), & -2 \leq x < -1, \\ \frac{1}{6}(2x + 3), & -1 \leq x < 1, \\ \frac{1}{6}(-x^2 + 4x + 2), & 1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

Η κατανομή εμφανίζεται στο Σχήμα 7.2.

Παράδειγμα 7.3. Έστω πως η διάρκεια ζωής X, σε χρόνια, μιας οθόνης υπολογιστή είναι μια συνεχής Τ.Μ. με πυκνότητα

$$f(x) = \begin{cases} 0, & x < 0, \\ e^{-x}, & x \geq 0. \end{cases}$$

Η γραφική της παράσταση δίνεται στο Σχήμα 7.4.

Παρατηρήστε κατ΄ αρχήν πως η $f(x)$ είναι παντού μη αρνητική, και επιπλέον

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{0} 0\, dx + \int_{0}^{\infty} e^{-x}\, dx = 0 + \lim_{t \to \infty} \int_{0}^{t} (-e^{-x})'\, dx = 1.$$

Στην πρώτη ισότητα λάβαμε υπ΄ όψιν ότι $f(x) = 0$ για $x < 0$. Πράγματι, λοιπόν, ικανοποιεί τις προϋποθέσεις για να είναι πυκνότητα.

Αν θέλουμε να υπολογίσουμε την πιθανότητα $P(1 \leq X \leq 2)$, έχουμε

$$P(1 \leq X \leq 2) = \int_{1}^{2} f(x)\, dx = \int_{1}^{2} e^{-x}\, dx = \big[-e^{-x}\big]_1^2 = e^{-1} - e^{-2}.$$

Αν θέλουμε να υπολογίσουμε την $P(-5 < X < 2)$, έχουμε

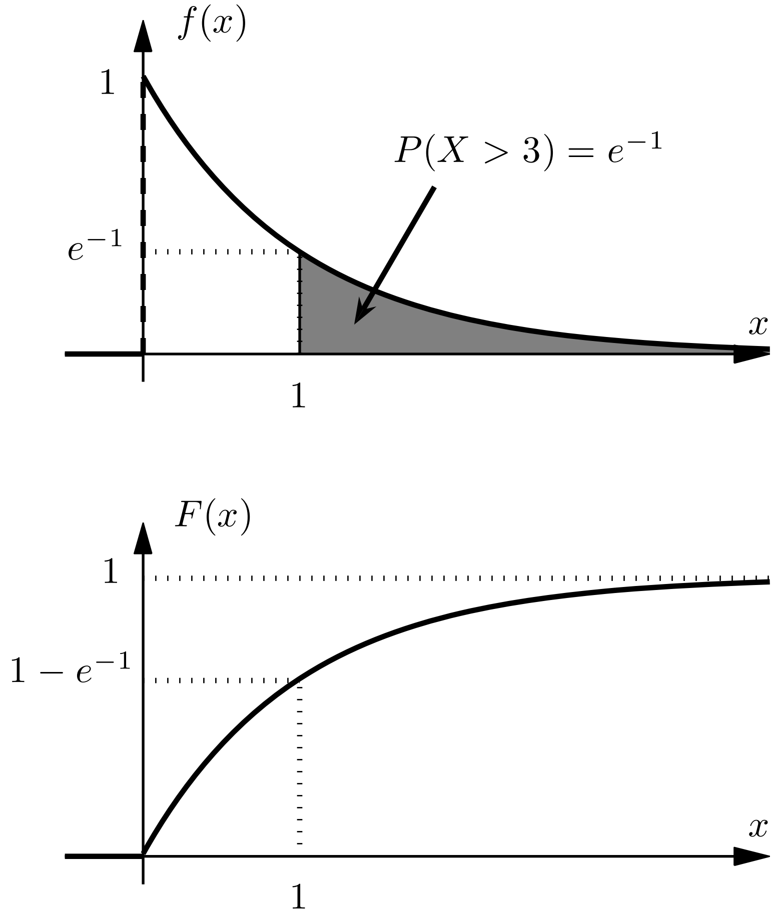
$$P(-5 < X < 2) = \int_{-5}^{2} f(x)\, dx = \int_{0}^{2} e^{-x}\, dx = \big[-e^{-x}\big]_0^2 = 1 - e^{-2}.$$

Για να υπολογίσουμε την κατανομή $F(x)$, παρατηρούμε κατ΄ αρχήν πως για $x < 0$ θα έχουμε:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt = 0,$$

αφού η ολοκληρωτέα συνάρτηση είναι 0 παντού στο διάστημα ολοκλήρωσης. Επιπλέον, για $x \geq 0$ έχουμε:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt = \int_{0}^{x} e^{-t}\, dt = \big[-e^{-t}\big]_0^x = 1 - e^{-x}.$$



Σχήμα 7.4: Γραφικές παραστάσεις της πυκνότητας $f(x)$ και της κατανομής $F(x)$ της Τ.Μ. X στο Παράδειγμα 7.3.

Συγκεντρωτικά,

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-x}, & x \geq 0. \end{cases}$$

Οι γραφικές παραστάσεις της πυκνότητας και της κατανομής δίνονται στο Σχήμα 7.4.

Το ότι γνωρίζουμε την συνάρτηση κατανομής μας διευκολύνει σημαντικά στον υπολογισμό πιθανοτήτων σχετικά με την X. Για παράδειγμα, η πιθανότητα $P(X > 1)$ μπορεί να υπολογιστεί απευθείας από την πυκνότητα ως

$$P(X > 1) = P(1 < X < \infty) = \int_1^\infty f(x)\,dx = \int_1^\infty e^{-x}\,dx = \big[-e^{-x}\big]_1^\infty = e^{-1},$$

ή, εναλλακτικά (και ευκολότερα), μέσω της συνάρτησης κατανομής:

$$P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = 1 - (1 - e^{-1}) = e^{-1}.$$

Στην πρώτη ισότητα χρησιμοποιήσαμε το γεγονός ότι το ενδεχόμενο $\{X > 1\}$ είναι συμπλήρωμα του $\{X \leq 1\}$. Παρομοίως μπορούμε να υπολογίσουμε και δεσμευμένες

πιθανότητες για την X. Για παράδειγμα:

$$\begin{aligned} P(X<4|X>3) &= \frac{P(3<X<4)}{P(X>3)} = \frac{P(3<X\leq 4)}{1-P(X\leq 3)} = \frac{P(X\leq 4)-P(X\leq 3)}{1-P(X\leq 3)} \\ &= \frac{F(4)-F(3)}{1-F(3)} = \frac{(1-e^{-4})-(1-e^{-3})}{1-(1-e^{-3})} = 1-e^{-1} \simeq 0.6321. \end{aligned}$$

Στην δεύτερη ισότητα χρησιμοποιήσαμε το ότι οι δύο πιθανότητες των δύο αριθμητών δίνονται από το ίδιο ολοκλήρωμα. Στην τρίτη, χρησιμοποιήσαμε την γνωστή ιδιότητα $P(A\cap B') = P(A) - P(A\cap B)$, όπου $A = \{X \leq 4\}$, $B = \{X \leq 3\}$.

Λήμμα 7.1. (Βασικές Ιδιότητες Πυκνότητας) *Έστω συνεχής Τ.Μ. X με πυκνότητα $f(x)$.*

1. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1.$

2. $$\int_a^b f(x)\,dx$$
   $$= P(a\leq X\leq b) = P(a<X\leq b) = P(a\leq X<b) = P(a<X<b).$$

   για κάθε $a, b \in \mathbb{R}$, με $a < b$.

3. $P(X = a) = 0.$

4. $P(X\geq a) = P(X>a) = \displaystyle\int_a^{\infty} f(x)\,dx.$

5. $P(X\leq a) = P(X<a) = \displaystyle\int_{-\infty}^{a} f(x)\,dx.$

Απόδειξη. Όλες προκύπτουν άμεσα από τον Ορισμό 7.1. □

Λήμμα 7.2. (Βασικές Ιδιότητες Κατανομής) *Έστω συνεχής Τ.Μ. X με πυκνότητα $f(x)$ και κατανομή $F(x)$.*

1. *Η $F(x)$ είναι συνεχής.*

2. *Όπου η $f(x)$ είναι συνεχής, ισχύει ότι*

$$F'(x) = f(x).$$

3. $\lim_{x\to\infty} F(x) = 1$ *και* $\lim_{x\to-\infty} F(x) = 0$.

4. $P(X \le a) = P(X < a) = F(a)$.

5. $P(X \ge a) = P(X > a) = 1 - F(a)$.

6. *Έστω* $a < b$.

$$\begin{aligned} &F(b) - F(a) \\ &\quad = P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b). \end{aligned}$$

Απόδειξη. 1. Προκύπτει άμεσα από το ότι η κατανομή είναι ολοκλήρωμα με το x να εμφανίζεται ως όριο ολοκλήρωσης.

2. Προκύπτει από το Θεμελιώδες Θεώρημα του Ολοκληρωτικού Λογισμού.

3. Η απόδειξη είναι μεγάλη και παραλείπεται.

4. Εξ ορισμού $P(X \le a) = F(a)$, ενώ έχουμε $P(X < a) = P(X \le a)$ με χρήση του Λήμματος 7.1.

5. Παρατηρήστε πως $P(X > a) = 1 - P(\{X > a\}') = 1 - P(X \le a) = 1 - F(a)$, ενώ από το Λήμμα 7.1 έχουμε ότι $P(X \ge a) = P(X > a)$.

6. Έχουμε

$$\begin{aligned} F(b) = P(X \le b) &= P(\{-\infty < X \le a\} \cup \{a < X \le b\}) \\ &= P(-\infty < X \le a) + P(a < X \le b) = F(a) + P(a < X \le b) \\ &\Rightarrow P(a < X \le b) = F(b) - F(a). \end{aligned}$$

Οι υπόλοιπες ισότητες προκύπτουν παρατηρώντας ότι η πιθανότητα να πάρει η X οποιαδήποτε συγκεκριμένη τιμή είναι 0. Για παράδειγμα:

$$P(a \le X \le b) = P(a < X \le b) + P(X = a) = F(b) - F(a) + 0 = F(b) - F(a).$$

□

7.2 Παραδείγματα

Παρατήρηση: Πολύ συχνά ο προσδιορισμός μιας «κατάλληλης» πυκνότητας, δηλαδή μιας πυκνότητας που ανταποκρίνεται καλά στο φυσικό πρόβλημα, είναι ουσιώδες μέρος του μαθηματικού προβλήματος. Δείτε το επόμενο παράδειγμα.

Παράδειγμα 7.4. Έστω, για παράδειγμα, πως ο χρόνος X (σε δευτερόλεπτα) που απαιτείται για την εκκίνηση της λειτουργίας ενός δικτύου είναι πάντοτε μεταξύ 10 και 20 δευτερολέπτων, και κατά τα άλλα είναι «εντελώς τυχαίος», δεν εμφανίζεται δηλαδή η τάση κάποιο εύρος τιμών να εμφανίζεται πιο συχνά από κάποιο άλλο. Για να περιγράψουμε το χρόνο X ως μια Τ.Μ. θα θέλαμε η πιθανότητα του να πάρει οποιαδήποτε τιμή στο διάστημα $[10, 20]$ να είναι κατά κάποιο τρόπο ομοιόμορφη. Π.χ., θα θέλαμε η πιθανότητα το X να είναι μεταξύ 10 και 11 δευτερολέπτων να είναι η ίδια με την πιθανότητα να έχουμε $19 \leq X \leq 20$. Επιπλέον, θα θέλαμε να ικανοποιεί την

$$P(10 \leq X \leq 15) = P(15 \leq X \leq 20) = \frac{1}{2}. \tag{7.3}$$

Η άνω σχέση σημαίνει πως θα είναι το ίδιο πιθανό η εκκίνηση να γίνει τα πρώτα 5 ή τα τελευταία 5 δευτερόλεπτα.

Τα άνω μπορούν να επιτευχθούν αν ορίσουμε την πυκνότητα $f(x)$ ως μια *σταθερά* $c > 0$, για $x \in [10, 20]$:

$$f(x) = \begin{cases} c, & x \in [10, 20], \\ 0, & x \notin [10, 20]. \end{cases}$$

Για να προσδιορίσουμε την τιμή της σταθεράς c παρατηρούμε πως θα πρέπει το ολοκλήρωμα της πυκνότητας σε όλο το $\mathbb{R}$ να είναι μονάδα:

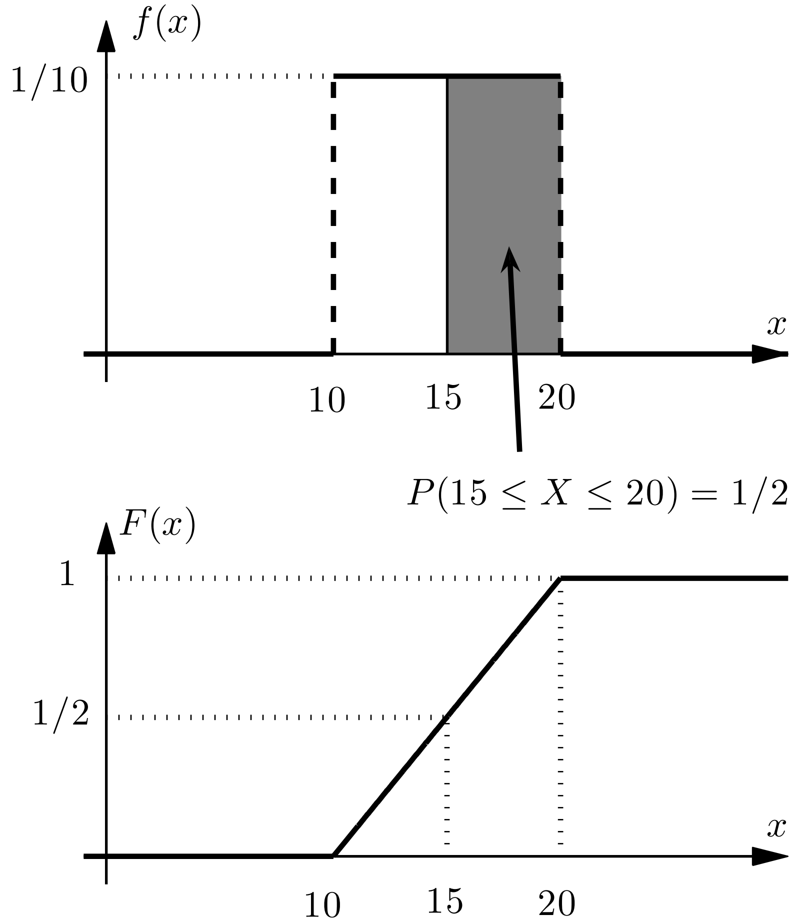
$$1 = \int_{-\infty}^{\infty} f(x)\, dx = \int_{10}^{20} c\, dx = \big[cx\big]_{10}^{20} = 20c - 10c = 10c \Rightarrow c = \frac{1}{10}.$$

Στην δεύτερη ισότητα αλλάξαμε τα όρια λαμβάνοντας υπ΄ όψιν ότι η πυκνότητα είναι μηδενική εκτός του $[10, 20]$.

Για να ελέγξουμε αν αυτή η Τ.Μ. πράγματι έχει ιδιότητες που ανταποκρίνονται στις απαιτήσεις μας ως προς την ποσότητα την οποία θέλουμε να περιγράψουμε, παρατηρούμε πως

$$P(15 \leq X \leq 20) = \int_{15}^{20} f(x)\, dx = \int_{15}^{20} \frac{1}{10}\, dx = \frac{20 - 15}{10} = \frac{1}{2},$$

και παρομοίως βρίσκουμε πως και η πιθανότητα $P(10 \leq X \leq 15)$ ισούται με $\frac{1}{2}$. Συνεπώς η σχέση (7.3), την οποία διατυπώσαμε διαισθητικά, επαληθεύεται και μαθηματικά.



Σχήμα 7.5: Γραφικές παραστάσεις της πυκνότητας $f(x)$ και της κατανομής $F(x)$ της Τ.Μ. X στο Παράδειγμα 7.4.

Η κατανομή $F(x)$ της X υπολογίζεται εύκολα με χρήση της (7.2). Παρατηρούμε καταρχήν πως αν $x \in [10, 20]$, τότε

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \int_{10}^{x} \frac{1}{10}\,dt = \frac{x-10}{10}.$$

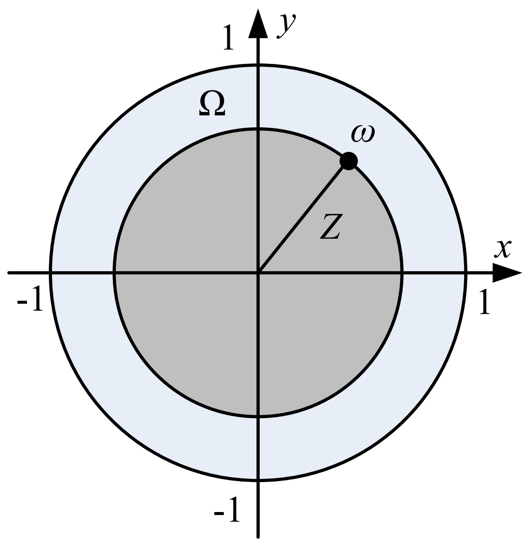
Αν $x > 20$, έχουμε

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \int_{10}^{20} \frac{1}{10}\,dt = 1,$$

αφού η ολοκληρωτέα συνάρτηση για $x > 20$ γίνεται μηδενική. Τέλος, για $x < 10$, έχουμε

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = 0,$$

αφού η ολοκληρωτέα συνάρτηση είναι παντού μηδέν μεταξύ των ορίων ολοκλήρωσης.



Σχήμα 7.6: Η Τ.Μ. Z του Παραδείγματος 7.5.

Συνοψίζοντας:

$$F(x) = \begin{cases} 0, & x < 10, \\ \frac{x-10}{10}, & x \in [10, 20], \\ 1, & x > 20. \end{cases}$$

Οι γραφικές παραστάσεις της πυκνότητας και της κατανομής δίνονται στο Σχήμα 7.5.

Παρατήρηση: Σε πολλές περιπτώσεις ο υπολογισμός της πυκνότητας ή/και της κατανομής είναι ένα πρόβλημα πιθανοτήτων από μόνος του. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 7.5. *(Τυχαίο σημείο σε κύκλο)* Επιλέγουμε ένα τυχαίο σημείο ω εντός του κλειστού μοναδιαίου κυκλικού δίσκου Ω. Έστω Z η απόσταση του σημείου από την αρχή των αξόνων. Το σύνολο τιμών της Z είναι το $[0, 1]$. Δείτε το Σχήμα 7.6. Έστω πως το σημείο ω επιλέγεται ομοιόμορφα, έτσι ώστε η πιθανότητα να βρεθεί σε κάποιο υποσύνολο K του μοναδιαίου κύκλου να είναι ίση με $|K|/\pi$, όπου $|K|$ είναι το εμβαδόν του K, και ανεξάρτητη από οτιδήποτε άλλο. Τότε προκύπτει πως

$$F_Z(z) = P(Z \leq z) = \begin{cases} 0, & z < 0, \\ \pi z^2/\pi = z^2, & 0 \leq z \leq 1, \\ 1, & z > 1. \end{cases}$$

Παραγωγίζοντας,

$$f_Z(z) = F'_Z(z) = \begin{cases} 0, & z < 0, \\ 2z, & 0 \leq z \leq 1, \\ 1, & z > 1. \end{cases}$$

Παρατήρηση: Πολύ συχνά η πυκνότητα ή/και η κατανομή δίνονται μέσω εκφράσεων που περιλαμβάνουν άγνωστες παραμέτρους. Δείτε το ακόλουθο παράδειγμα:

Παράδειγμα 7.6. *(Πυκνότητα με άγνωστη παράμετρο)* Έστω ότι ο χρόνος (σε δευτερόλεπτα) X μέχρι την εμφάνιση του πρώτου πακέτου σε έναν router είναι μια συνεχής Τ.Μ. με πυκνότητα

$$f(x) = \begin{cases} C, & 0 \leq x \leq 5, \\ \frac{C}{5}(10 - x), & 5 < x \leq 10, \\ 0, & x \notin [0, 10]. \end{cases}$$

Θα απαντήσουμε τα ακόλουθα ερωτήματα:

1. Ποια είναι η τιμή του C;
2. Πόση είναι η πιθανότητα $P(X < 5)$;
3. Πόση είναι η πιθανότητα $P\left(X \leq \frac{1}{2} | X < 5\right)$;

Έχουμε, κατά περίπτωση,

1. Για να είναι η $f(x)$ πυκνότητα πρέπει:

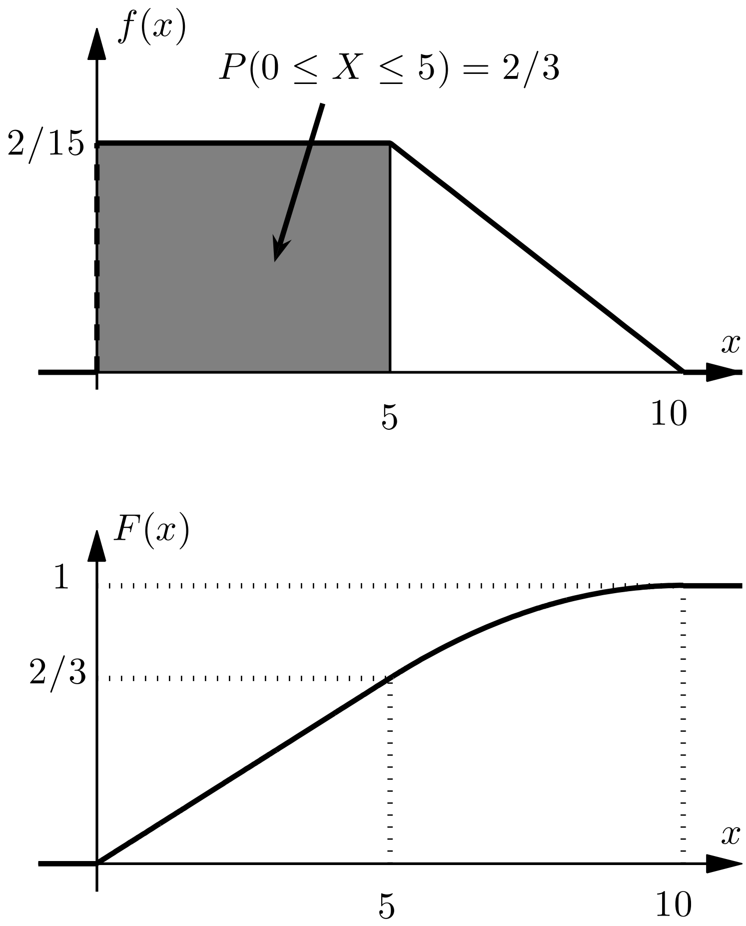
$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \Rightarrow \int_0^5 C\, dx + \int_5^{10} \frac{C}{5}(10 - x)\, dx = 1 \Rightarrow 5C + \frac{C}{5}\Big[10x - \frac{x^2}{2}\Big]_5^{10} = 1,$$

και άρα

$$5C + \frac{C}{5}(100 - \frac{100}{2} - 50 + \frac{25}{2}) = 1 \Rightarrow \frac{15C}{2} = 1 \Rightarrow C = \frac{2}{15}.$$

Η πυκνότητα $f(x)$, καθώς και η κατανομή $F(x)$, που εύκολα μπορεί να προσδιοριστεί ολοκληρώνοντας την $f(x)$, εμφανίζονται στο Σχήμα 7.7.

2. $P(X < 5) = \int_0^5 f(x)\, dx = \int_0^5 \frac{2}{15}\, dx = \frac{2}{15} \times 5 = \frac{2}{3}.$



Σχήμα 7.7: Παράδειγμα 7.6.

3.

$$P\left(X \leq \frac{1}{2}|X<5\right) = \frac{P(X \leq \frac{1}{2},\ X<5)}{P(X<5)} = \frac{P(X \leq \frac{1}{2})}{P(X<5)}$$
$$= \frac{\int_0^{\frac{1}{2}} \frac{2}{15}\, dx}{\frac{2}{3}} = \frac{\frac{1}{2} \times \frac{2}{15}}{\frac{2}{3}} = \frac{1}{10}.$$

Σημείωση: Και στα τρία υποερωτήματα, αντί για τον υπολογισμό των ολοκληρωμάτων όπως πιο πάνω θα μπορούσαμε να έχουμε απλά υπολογίσει το αντίστοιχο εμβαδόν από το διάγραμμα. Για παράδειγμα, στο δεύτερο σκέλος έχουμε ότι η ζητούμενη πιθανότητα $P(X<5)$ είναι απλά το εμβαδόν του ορθογωνίου με βάση το τμήμα του άξονα x από το 0 ως το 5 και ύψος το τμήμα του άξονα y από το μηδέν ως το $C = 2/15$. Άρα η πιθανότητα $P(X<5)$ ισούται με $(5-0) \times (2/15-0) = 2/3$.

7.3 Μέση Τιμή και Διασπορά

Η μέση τιμή και η διασπορά για μια συνεχή Τ.Μ. X ορίζονται κατά τρόπο ανάλογο με εκείνον που είδαμε στην περίπτωση διακριτών Τ.Μ.

Ορισμός 7.3. (Μέση τιμή) *Έστω συνεχής Τ.Μ. X με πυκνότητα $f(x)$. Ορίζουμε ως τη μέση τιμή ή αναμενόμενη τιμή ή προσδοκώμενη τιμή της X το ολοκλήρωμα*

$$E(X) \triangleq \int_{-\infty}^{\infty} x f(x)\, dx.$$

Παρατηρήσεις

1. Συχνά χρησιμοποιείται το σύμβολο μ για την μέση τιμή μιας Τ.Μ.

2. Οι ορισμοί των μέσων τιμών στην διακριτή και την συνεχή περίπτωση είναι εντελώς αντίστοιχοι. Παρατηρήστε ότι στη διακριτή περίπτωση

$$E(X) = \sum_{x \in S} x p(x).$$

Άρα, τον ρόλο του αθροίσματος τον παίζει το ολοκλήρωμα, και στη θέση της μάζας βρίσκεται η ποσότητα $f(x)dx$, που όπως ήδη έχουμε δει εκφράζει προσεγγιστικά την πιθανότητα να λάβει η X τιμές στο διάστημα $[x - dx/2, x + dx/2]$. Η αναλογία γίνεται πιο ξεκάθαρη αν γράψουμε τον ορισμό του άνω ολοκληρώματος Riemann για την ειδική περίπτωση όπου το διάστημα ολοκλήρωσης είναι το $[a, b]$:

$$E(X) = \int_a^b x f(x)\, dx \triangleq \lim_{\Delta x_i \to 0,\ n \to \infty} \sum_{i=1}^{n} x_i f(x_i) \Delta x_i.$$

Υπενθυμίζουμε ότι στον άνω ορισμό έχουμε διαμερίσει το $[a, b]$ σε ένα σύνολο n διαστημάτων των οποίων το πλάτος Δx_i τείνει στο 0, ενώ το x_i έχει επιλεγεί από το i-οστό διάστημα.

3. Διαισθητικά, όπως και στη διακριτή περίπτωση, περιμένουμε ότι ο μέσος όρος πολλών Τ.Μ. κατανεμημένων όπως η Τ.Μ. X θα είναι κοντά στη μέση τιμή $E(X)$. Διαισθητικά, επίσης, περιμένουμε ότι οι τιμές διαδοχικών Τ.Μ. κατανεμημένων όπως η X «τείνουν να κυμαίνονται» γύρω από την $E(X)$.

Παράδειγμα 7.7. Ένας μεγιστάνας θα ζήσει ακόμα ένα χρονικό διάστημα X (σε έτη) που έχει την ακόλουθη πυκνότητα:

$$f(x) = \begin{cases} 2 - 2x, & x \in [0, 1], \\ 0, & \text{αλλού.} \end{cases}$$

Το μέσο διάστημα που του μένει να ζήσει είναι

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x)\, dx = \int_0^1 x(2-2x)\, dx \\ &= \int_0^1 (x^2)'\, dx - \int_0^1 \left(\frac{2x^3}{3}\right)' dx = 1 - \frac{2}{3} = \frac{1}{3}. \end{aligned}$$

Λήμμα 7.3. (Μέση τιμή συνάρτησης Τ.Μ.) *Έστω X συνεχής Τ.Μ. με πυκνότητα $f_X(x)$.*

1. *Έστω $Y = g(X)$ μια νέα Τ.Μ., συνάρτηση της προηγούμενης, με μέση τιμή $E(Y)$. Θα ισχύει:*

$$E(Y) = E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx.$$

2. *Αν $Y = aX + b$, όπου $a, b \in \mathbb{R}$, έχουμε $E(aX+b) = aE(X) + b$.*

3. *Πιο γενικά, αν $Y = a_1 g_1(X) + a_2 g_2(X) + \cdots + a_n g_n(X) = \sum_{i=1}^n a_i g_i(X)$, τότε*

$$E\left(\sum_{i=1}^n a_i g_i(X)\right) = \sum_{i=1}^n a_i E(g_i(X)).$$

Απόδειξη. 1. Είναι αρκετά σύνθετη και παραλείπεται.

2. Χρησιμοποιώντας το προηγούμενο σκέλος, παρατηρούμε πως:

$$\begin{aligned} E(aX+b) &= \int_{-\infty}^{\infty} (ax+b) f_X(x)\, dx \\ &= a\int_{-\infty}^{\infty} x f_X(x)\, dx + b\int_{-\infty}^{\infty} f_X(x)\, dx = aE(X) + b. \end{aligned}$$

3. Πιο γενικά, πάλι με χρήση του πρώτου σκέλους, έχουμε:

$$\begin{aligned} E\left(\sum_{i=1}^n a_i g_i(X)\right) &= \int_{-\infty}^{\infty} \sum_{i=1}^n a_i g_i(x_i) f_X(x)\, dx \\ &= \sum_{i=1}^n a_i \int_{-\infty}^{\infty} g_i(x_i) f_X(x)\, dx = \sum_{i=1}^n a_i E(g_i(X)). \end{aligned}$$

□

Παρατήρηση: Το άνω λήμμα ισχύει ανεξάρτητα από το αν η Y είναι διακριτή, ή συνεχής, ή τίποτα από τα δύο! (Παρατηρήστε ότι, για την τελευταία περίπτωση, δεν έχουμε ορίσει την μέση τιμή.)

Παράδειγμα 7.8. Επιστρέφοντας στο Παράδειγμα 7.7, έστω πως ο μεγιστάνας παντρεύεται μια νεαρή ηθοποιό, και το προγαμιαίο συμβόλαιο προβλέπει πως κατά το θάνατο του μεγιστάνα η ηθοποιός θα εισπράξει κληρονομιά $Y = 10X + 2$ εκατομμύρια δολάρια. Η μέση τιμή της κληρονομιάς θα ισούται με

$$E(Y) = E(10X + 2) = 10E(X) + 2 = 10 \times \frac{1}{3} + 2 = \frac{16}{3}.$$

Επιπλέον,

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 x^2(2-2x)\,dx \\ &= \int_0^1 \left(\frac{2}{3}x^3\right)' dx - \int_0^1 \left(\frac{x^4}{2}\right)' dx = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

Ορισμός 7.4. (Διασπορά) *Έστω συνεχής Τ.Μ. X με μέση τιμή $E(X)$. Η διασπορά της ορίζεται ως*

$$\mathrm{VAR}(X) = E\left[(X - E(X))^2\right].$$

Η τυπική απόκλιση της X ορίζεται ως $\sqrt{\mathrm{VAR}(X)}$.

Παρατηρήσεις

1. Συνεπώς, η διασπορά ορίζεται ανάλογα με την περίπτωση των διακριτών Τ.Μ., και εκφράζει το πόσο «απλωμένη» είναι η Τ.Μ. γύρω από τη μέση τιμή της.

2. Όπως και στη διακριτή περίπτωση, χρησιμοποιούμε επίσης το σύμβολο σ^2 για τη διασπορά, και συνεπώς το σ για την τυπική απόκλιση.

Λήμμα 7.4. (Ιδιότητες διασποράς) *Έστω Τ.Μ. X με μέση τιμή $E(X)$ και διασπορά $\mathrm{VAR}(X)$. Ισχύουν τα ακόλουθα:*

1. $\mathrm{VAR}(X) = E(X^2) - (E(X))^2$.

2. $\mathrm{VAR}(aX + b) = a^2\mathrm{VAR}(X)$.

Απόδειξη. Η απόδειξη είναι εντελώς ανάλογη με αυτή της διακριτής περίπτωσης. □

Παράδειγμα 7.9. Επιστρέφοντας στα Παραδείγματα 7.7, 7.8, έχουμε

$$\mathrm{VAR}(X) = E(X^2) - (E(X))^2 = \frac{1}{6} - \left(\frac{1}{3}\right)^2 = \frac{1}{18}.$$

Βεβαιωθείτε ότι μπορείτε να φτάσετε στο ίδιο αποτέλεσμα χρησιμοποιώντας τον ορισμό της διασποράς.

Παράδειγμα 7.10. *(Κανονικοποίηση)* Αν η X είναι Τ.Μ. με μέση τιμή $E(X) = \mu$ και διασπορά $\mathrm{VAR}(X) = \sigma^2$, τότε η Τ.Μ. $Y = \frac{X-\mu}{\sigma}$ έχει μέση τιμή $E(Y) = 0$ και διασπορά $\mathrm{VAR}(Y) = 1$. Η απόδειξη είναι ακριβώς όπως και στη διακριτή περίπτωση.

Παράδειγμα 7.11. Θα υπολογίσουμε τη μέση τιμή της Τ.Μ. X του Παραδείγματος 7.3. Θα δείξουμε δύο τρόπους υπολογισμού του καταχρηστικού ολοκληρώματος που προκύπτει.

Ο πρώτος ακολουθεί αυστηρά τον ορισμό των καταχρηστικών ολοκληρωμάτων:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x)\,dx = \lim_{t\to-\infty}\int_t^0 f(x)\,dx + \lim_{t\to\infty}\int_0^t f(x)\,dx \\
&= 0 + \lim_{t\to\infty}\int_0^t x e^{-x}\,dx = \lim_{t\to\infty}\left[\left[-xe^{-x}\right]_0^t + \int_0^t e^{-x}\,dt\right] \\
&= \lim_{t\to\infty}\left[\left[-xe^{-x}\right]_0^t + \left[-e^{-x}\right]_0^t\right] = \lim_{t\to\infty}\left[-xe^{-x}\right]_0^t + \lim_{t\to\infty}\left[-e^{-x}\right]_0^t \\
&= \lim_{t\to\infty}\left[-te^{-t} + 0\right] + \lim_{t\to\infty}\left[1 - e^{-t}\right] = \lim_{t\to\infty}\frac{t}{e^t} + 1 = \lim_{t\to\infty}\frac{1}{e^t} + 1 = 1.
\end{aligned}$$

Στην τέταρτη ισότητα θέσαμε $u = x$ και $dv = e^{-x}dx$, έτσι ώστε $du = dx$ και $v = -e^{-x}$, και εφαρμόσαμε ολοκλήρωση κατά παράγοντες. Στην προτελευταία εξίσωση κάναμε χρήση του κανόνα του L'Hôspital.

Ο δεύτερος τρόπος βασίζεται στην ίδια θεωρία, αλλά είναι πιο συνοπτικός:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^0 f(x)\,dx + \int_0^{\infty} f(x)\,dx \\
&= 0 + \int_0^{\infty} x e^{-x}\,dx = \left[-xe^{-x}\right]_0^{\infty} + \int_0^{\infty} e^{-x}\,dt \\
&= \left[-xe^{-x}\right]_0^{\infty} + \left[-e^{-x}\right]_0^{\infty} = \lim_{t\to\infty}\left[-xe^{-x}\right]_0^t + \lim_{t\to\infty}\left[-e^{-x}\right]_0^t \\
&= \cdots = 1.
\end{aligned}$$

Στην τελευταία γραμμή συνεχίζουμε τις πράξεις όπως με τον πρώτο τρόπο. Παρατηρήστε πως αντιμετωπίσαμε το ∞ όποτε εμφανίστηκε ως όριο ολοκληρώματος ως ένα κοινό αριθμό, μέχρι που χρειάστηκε να το απλοποιήσουμε, οπότε και καταφύγαμε στον υπολογισμό ενός ορίου.

Με τον ίδιο συνοπτικό τρόπο υπολογίζουμε και την μέση τιμή του X^2:

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2\, f(x)\, dx = \int_0^{\infty} x^2\, e^{-x}\, dx = \big[-x^2 e^{-x}\big]_0^{\infty} + \int_0^{\infty} 2xe^{-x}\, dx \\
&= \big[-x^2 e^{-x}\big]_0^{\infty} + 2E(X) = \lim_{t\to\infty} \left[-x^2 e^{-x}\right]_0^{t} + 2 \\
&= \lim_{t\to\infty} \left[-t^2 e^{-t} + 0\right] + 2 = -\lim_{t\to\infty} \frac{t^2}{e^t} + 2 = -\lim_{t\to\infty} \frac{2t}{e^t} + 2 = -\lim_{t\to\infty} \frac{2}{e^t} + 2 = 2.
\end{aligned}$$

Στην τρίτη ισότητα θέσαμε $u = x^2$ και $dv = e^{-x}dx$, έτσι ώστε $du = 2xdx$ και $v = -e^{-x}$, και κατόπιν εφαρμόσαμε ολοκλήρωση κατά παράγοντες. Στην τέταρτη ισότητα χρησιμοποιήσαμε το άνω ολοκλήρωμα. Στο τέλος, κάναμε διπλή χρήση του κανόνα του L'Hôspital. Συνεπώς,

$$\mathrm{VAR}(X) = E(X^2) - (E(X))^2 = 2 - 1^2 = 1.$$

Παρατήρηση: Στα επόμενα, θα χρησιμοποιούμε παγίως τον συνοπτικό τρόπο που περιγράφηκε στο άνω παράδειγμα για να υπολογίζουμε καταχρηστικά ολοκληρώματα. Είναι όμως σημαντικό να έχουμε κατανοήσει πλήρως και τους δύο. Σε περίπτωση σύγχυσης, συνίσταται η χρήση του αναλυτικού τρόπου, που ακολουθεί πιο πιστά τον ορισμό.

Παράδειγμα 7.12. Η μέση τιμή της Τ.Μ. X του Παραδείγματος 7.4 είναι:

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{10}^{20} x \frac{1}{10}\, dx = \frac{1}{20} \int_{10}^{20} (x^2)'\, dx = \frac{1}{20}(20^2 - 10^2) = 15.$$

Το αποτέλεσμα ήταν αναμενόμενο.

Έστω τώρα πως η X εκφράζει την απόσταση που διανύει μια βολή κανονιού που έχει ένα στόχο που βρίσκεται σε απόσταση 15 χιλιομέτρων. Συνεπώς, ο χειριστής του κανονιού πετυχαίνει το στόχο αν $X = 15$, σε περίπτωση που η βολή είναι $X = 20$ το σφάλμα είναι 5 χιλιόμετρα, κ.ο.κ. Έστω πως ο διοικητής του χειριστή του επιβάλλει ποινή $Y = (X - 15)^2$ μέρες. Με χρήση του Λήμματος 7.3,

$$\begin{aligned}
E(Y) = E\left((X-15)^2\right) &= \int_{-\infty}^{\infty} (x-15)^2 f(x)\, dx = \int_{10}^{20} (x-15)^2 \frac{1}{10}\, dx \\
&= \frac{1}{30} \int_{10}^{20} \left[(x-15)^3\right]'\, dx = \frac{1}{30} \left[(x-15)^3\right]_{10}^{20} = \frac{25}{3}.
\end{aligned}$$

Με παρόμοιο τρόπο υπολογίζουμε και την μέση τιμή του X^2:

$$E(X^2) = \int_{-\infty}^{\infty} x^2\, f(x)\, dx = \int_{10}^{20} x^2 \frac{1}{10}\, dx = \frac{1}{30} \int_{10}^{20} (x^3)'\, dx = \frac{1}{30} \left[x^3\right]_{10}^{20} = \frac{700}{3}.$$

Συνεπώς,

$$\mathrm{VAR}(X) = E(X^2) - [E(X)]^2 = \frac{700}{3} - 15^2 = \frac{25}{3}.$$

7.4 Παραδείγματα

Παράδειγμα 7.13. Έστω μια συνεχής Τ.Μ. X με πυκνότητα

$$f(x) = \begin{cases} cx^2, & x \in [0,1], \\ 0, & x \notin [0,1]. \end{cases}$$

Η τιμή της c μπορεί να υπολογιστεί από την απαίτηση το ολοκλήρωμά της σε όλο το $\mathbb{R}$ να ισούται με τη μονάδα:

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 cx^2\,dx = \left[\frac{cx^3}{3}\right]_0^1 = \frac{c}{3},$$

και συνεπώς έχουμε $c = 3$. Η πυκνότητα έχει σχεδιαστεί στο Σχήμα 7.8.

Έχοντας την πυκνότητα, μπορούμε να υπολογίσουμε την πιθανότητα το X να βρεθεί σε οποιοδήποτε διάστημα. Για παράδειγμα:

$$P\left(\frac{1}{2} \leq X \leq 1\right) = \int_{1/2}^1 3x^2\,dx = \left[x^3\right]_{1/2}^1 = \frac{7}{8}.$$

Το εμβαδόν εμφανίζεται σκιασμένο στο Σχήμα 7.8. Ένας παρόμοιος υπολογισμός μας δίνει και την συνάρτηση κατανομής της X: προφανώς έχουμε $F(x) = P(X \leq x) = 0$ για $x < 0$, και $F(x) = P(X \leq x) = 1$ όταν $x > 1$, ενώ για $x \in [0,1]$ έχουμε

$$F(x) = P(X \leq x) = \int_0^x 3y^2\,dy = x^3,$$

οπότε

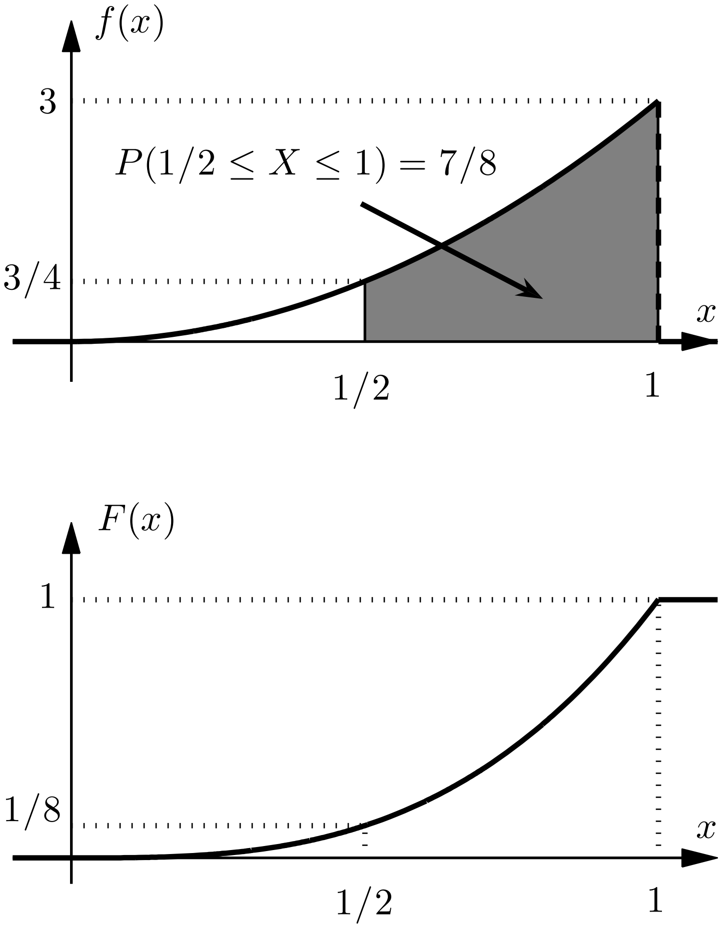
$$F(x) = \begin{cases} 0, & x < 0, \\ x^3, & x \in [0,1], \\ 1, & x > 1. \end{cases}$$

Η κατανομή έχει σχεδιαστεί στο Σχήμα 7.8. Γνωρίζοντας πλέον την κατανομή, η πιθανότητα $P(\frac{1}{2} \leq X \leq 1)$ που έχουμε ήδη υπολογίσει θα μπορούσε εναλλακτικά να προκύψει ως εξής:

$$P\left(\frac{1}{2} \leq X \leq 1\right) = F(1) - F\left(\frac{1}{2}\right) = 1 - \left(\frac{1}{2}\right)^3 = \frac{7}{8}.$$

Επίσης,

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x\,f(x)\,dx = \int_0^1 3x^3\,dx, = \left[\frac{3x^4}{4}\right]_0^1 = \frac{3}{4}, \\ E(X^2) &= \int_{-\infty}^{\infty} x^2\,f(x)\,dx = \int_0^1 3x^4\,dx = \left[\frac{3x^5}{5}\right]_0^1 = \frac{3}{5}, \\ \mathrm{VAR}(X) &= E(X^2) - (E(X))^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80} = 0.0375. \end{aligned}$$



Σχήμα 7.8: Γραφικές παραστάσεις της πυκνότητας και της κατανομής της Τ.Μ. X του Παραδείγματος 7.13.

Παράδειγμα 7.14. Έστω πως η συνάρτηση κατανομής μιας Τ.Μ. X είναι

$$F(x) = \begin{cases} C - 4/x^2, & x \geq 2, \\ 0, & x < 2. \end{cases}$$

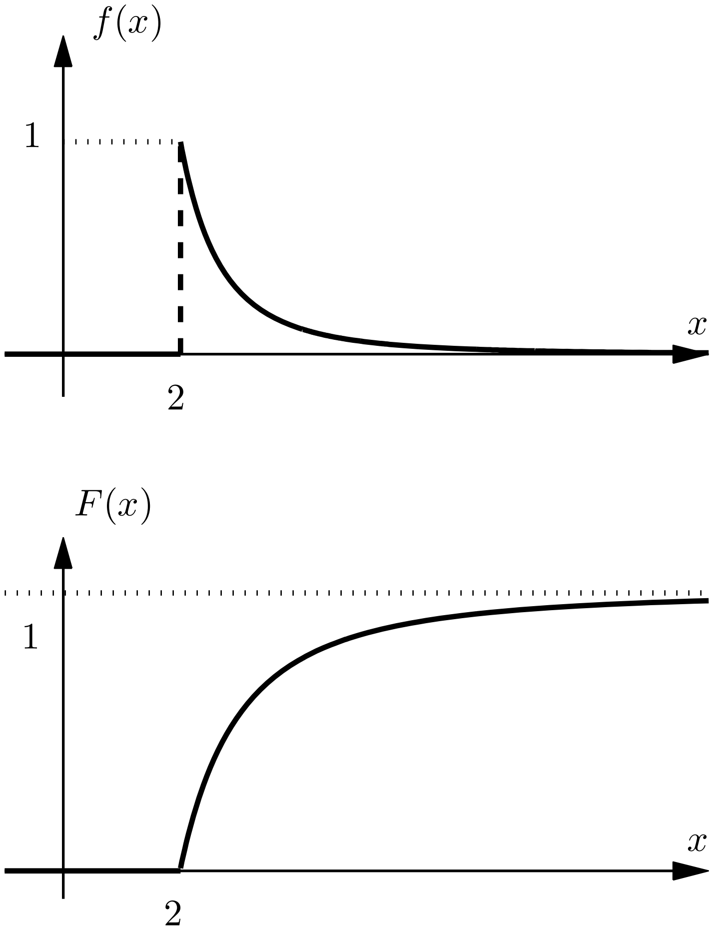
Θα βρούμε την τιμή του C, την πυκνότητα $f_X(x)$, και τη μέση τιμή $E(X)$.

Καταρχήν, για κάθε κατανομή ισχύει ότι $\lim_{x\to\infty} F(x) = 1$, άρα εδώ έχουμε:

$$\lim_{x\to\infty}\left(C - \frac{4}{x^2}\right) = 1 \Rightarrow C = 1.$$

Από το δεύτερο σκέλος του Λήμματος 7.2, η πυκνότητα ισούται με την παράγωγο της κατανομής, οπότε

$$f(x) = F'(x) = \begin{cases} 8x^{-3}, & x \geq 2, \\ 0, & x < 2. \end{cases}$$



Σχήμα 7.9: Άσκηση 7.14.

Η τιμή της πυκνότητας στο $x = 2$, όπου η δοσμένη κατανομή δεν είναι παραγωγίσιμη, επιλέχθηκε αυθαίρετα, αφού η τιμή της πυκνότητας σε ένα σημείο δεν μπορεί να επηρεάσει την τιμή των ολοκληρωμάτων της. Τέλος, η μέση τιμή ισούται με

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_{2}^{\infty} x \frac{8}{x^3}\, dx = \int_{2}^{\infty} \frac{8}{x^2}\, dx = \left[-\frac{8}{x}\right]_2^{\infty} = \frac{8}{2} = 4.$$

Η πυκνότητα $f_X(x)$ και η κατανομή $F_X(x)$ έχουν σχεδιαστεί στο Σχήμα 7.9.

7.5 Άλλα Είδη Τυχαίων Μεταβλητών

Ορισμός 7.5. (Τυχαία Μεταβλητή)

1. *Τυχαία Μεταβλητή (Τ.Μ.) καλείται κάθε συνάρτηση $X : \Omega \to \mathbb{R}$ όπου ο Ω είναι ένας δειγματικός χώρος.*

2. *Η συνάρτηση κατανομής πιθανότητας, ή κατανομή πιθανότητας, ή απλώς κατανομή της X είναι η συνάρτηση $F : \mathbb{R} \to [0, 1]$ η οποία ορίζεται ως*

$$F(x) = P(X \leq x) = P(\{\omega : X(\omega) \leq x\}), \quad \forall x \in \mathbb{R}.$$

Παρατηρήσεις

1. Στα προηγούμενα έχουμε δει τα δύο πιο χρήσιμα είδη τυχαίων μεταβλητών, τις συνεχείς και τις διακριτές.

2. Ένα άλλο ενδιαφέρον είδος είναι οι *μικτές Τ.Μ.* Μια Τ.Μ. καλείται μικτή αν με πιθανότητα p ισούται με μια συνεχή Τ.Μ., και με πιθανότητα $1 - p$ ισούται με μια διακριτή.

3. Υπάρχουν και Τ.Μ. που δεν ανήκουν σε κανένα από τα άνω είδη. Για όλες, όμως, υπάρχει η συνάρτηση κατανομής πιθανότητας.

Παράδειγμα 7.15. *(Μια μικτή Τ.Μ.)* Έστω πως ένα άτομο έχει 5 τσέπες. Το κινητό του βρίσκεται με πιθανότητα $1/20$ σε κάθε μια από τις 5 τσέπες, και με πιθανότητα $3/4$ κάπου στο σπίτι του. Τη χρονική στιγμή 0 το άτομο αρχίζει να ψάχνει τις τσέπες του διαδοχικά για να βρει το κινητό του, και ο χρόνος που απαιτείται για κάθε τσέπη είναι ακριβώς 1 δευτερόλεπτο. Αν το άτομο εξαντλήσει τις τσέπες, τότε ψάχνει το κινητό του στο σπίτι του, και ο χρόνος που απαιτείται σε αυτή την περίπτωση, μετρώντας από τη χρονική στιγμή 0, έχει πυκνότητα

$$f_Z(z) = \begin{cases} 0, & z < 5, \\ \frac{1}{10} e^{-(z-5)/10}, & z \geq 5. \end{cases}$$

Παρατηρήστε ότι $P(Z < 5) = 0$, αφού το άτομο αρχίζει να ψάχνει το κινητό στο σπίτι του μετά από 5 δευτερόλεπτα.

Παρατηρήστε ότι ο χρόνος T που απαιτείται για να βρει το άτομο το κινητό του δεν είναι ούτε διακριτή Τ.Μ. (αφού οι τιμές που μπορεί να πάρει έχουν μη αριθμήσιμο πλήθος), ούτε συνεχής, αφού έχει θετική πιθανότητα να πάρει τις τιμές $1, 2, 3, 4, 5$. Είναι όμως μια μικτή Τ.Μ., αφού με πιθανότητα $p = 1/4$ ισούται με μια διακριτή Τ.Μ.

X με

$$p_X(1) = p_X(2) = p_X(3) = p_X(4) = p_X(5) = \frac{1}{5},$$

ενώ με πιθανότητα $1 - p = 3/4$ ισούται με μια συνεχή Τ.Μ. Z.

Αν και δεν έχουμε δει αυστηρό ορισμό για τη μέση τιμή τέτοιων Τ.Μ., προφανώς, η μέση τιμή της συγκεκριμένης είναι

$$\begin{aligned} E(T) &= \frac{1}{4}E(X) + \frac{3}{4}E(Z) \\ &= \frac{1}{4} \times \left(1 \times \frac{1}{5} + 1 \times \frac{1}{5} + 2 \times \frac{1}{5} + 3 \times \frac{1}{5} + 4 \times \frac{1}{5} + 5 \times \frac{1}{5}\right) \\ &\quad + \frac{3}{4} \times \int_5^\infty \frac{z}{10} e^{-(z-5)/10}\, dz \\ &= \frac{1}{4} \times 3 + \frac{3}{4} \times 15 = 12. \end{aligned}$$

Παρατηρήστε ότι ο τύπος που χρησιμοποιήσαμε είναι μια μίξη των τύπων για την μέση τιμή που ισχύουν στην διακριτή και την συνεχή περίπτωση.

Παράδειγμα 7.16. *(Ακόμα μια μικτή Τ.Μ.)* Ο χρόνος T που χρειάζεται για να δούμε μια ιστοσελίδα είναι t_0, αν αυτή η ιστοσελίδα είναι ήδη τοπικά αποθηκευμένη στον υπολογιστή μας. Αλλιώς, ο χρόνος ακολουθεί την κατανομή

$$f(t) = \begin{cases} 0, & t < 0, \\ e^{-t}, & t \geq 0. \end{cases}$$

Επιπλέον, δίνεται ότι η πιθανότητα να υπάρχει η ιστοσελίδα τοπικά είναι p. Παρατηρήστε πως ο χρόνος T δεν είναι ούτε διακριτή Τ.Μ. (αφού μπορεί να πάρει μη αριθμήσιμο πλήθος τιμών), ούτε συνεχής, αφού έχει θετική πιθανότητα να πάρει την τιμή t_0. Είναι όμως μικτή.

Παρόμοια με το προηγούμενο παράδειγμα

$$E(T) = pt_0 + (1-p)\int_0^\infty te^{-t}\, dt = pt_0 + (1-p) \times 1 = pt_0 + 1 - p.$$

Κεφάλαιο 8

Συνήθεις Περιπτώσεις Συνεχών Τυχαίων Μεταβλητών

8.1 Ομοιόμορφη Κατανομή

Η πιο απλή περίπτωση μιας συνεχούς Τ.Μ. είναι εκείνη που παίρνει «ομοιόμορφα τυχαίες» τιμές σε κάποιο διάστημα $[a, b]$ στο $\mathbb{R}$, χωρίς να δείχνει προτίμηση σε κάποιο εύρος τιμών έναντι κάποιου άλλου του ίδιου μεγέθους.

Ορισμός 8.1. *Μια Τ.Μ. X έχει ομοιόμορφη κατανομή στο διάστημα $[a, b]$, για κάποια $a < b$, αν έχει πυκνότητα*

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

Για συντομία, γράφουμε $X \sim U[a, b]$.

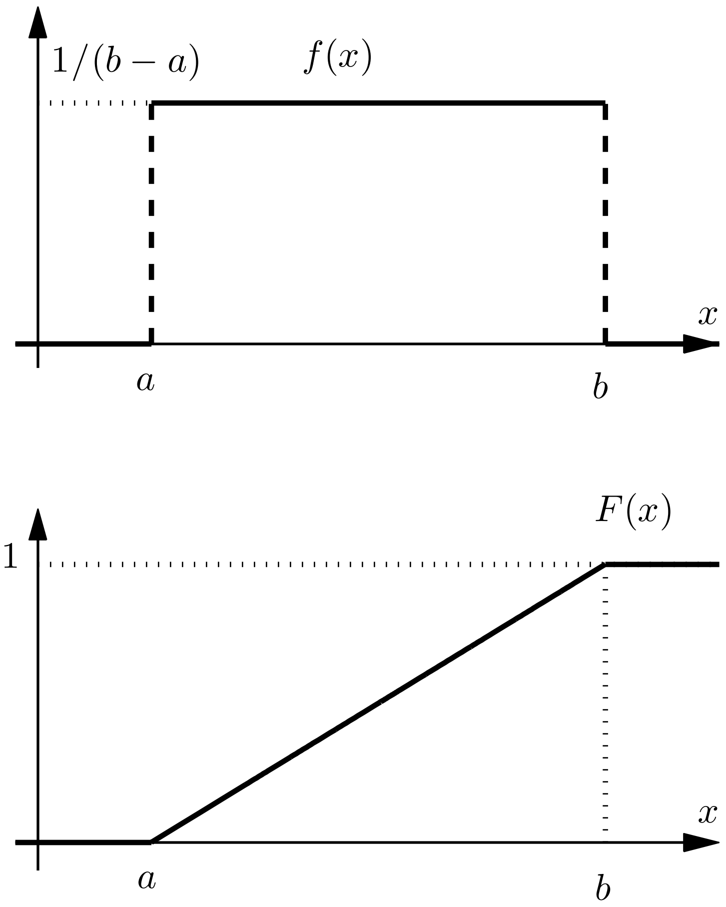
Λήμμα 8.1. (Ιδιότητες ομοιόμορφης κατανομής) *Έστω συνεχής Τ.Μ. X ομοιόμορφα κατανεμημένη στο διάστημα $[a, b]$. Ισχύουν τα ακόλουθα:*

1. $F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & x > b. \end{cases}$

2. $E(X) = \dfrac{a+b}{2}$.

3. $E(X^2) = \dfrac{a^2 + ab + b^2}{3}$.

4. $\mathrm{VAR}(X) = \dfrac{(b-a)^2}{12}$.



Σχήμα 8.1: Γραφική παράσταση της πυκνότητας $f(x)$ και της κατανομής $F(x)$ μιας Τ.Μ. X ομοιόμορφα κατανεμημένης στο διάστημα $[a, b]$.

Απόδειξη. 1. Προφανώς έχουμε

$$\begin{aligned} F(x) &= P(X \leq x) = 0, \quad \text{για } x < a, \\ F(x) &= P(X \leq x) = 1, \quad \text{για } x > b, \end{aligned}$$

ενώ για $x \in [a, b]$ βρίσκουμε

$$F(x) \;=\; P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \int_{a}^{x} \frac{1}{b-a}\,dt = \left[\frac{t}{b-a}\right]_a^x = \frac{x-a}{b-a}.$$

2. Εύκολα προκύπτει πως

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x\,f(x)\,dx = \int_{a}^{b} \frac{x}{b-a}\,dx = \left[\frac{x^2}{2(b-a)}\right]_a^b \\ &= \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}. \end{aligned}$$

3. Παρόμοια, έχουμε:

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_a^b \frac{x^2}{b-a}\,dx = \left[\frac{x^3}{3(b-a)}\right]_a^b \\ &= \frac{b^3-a^3}{3(b-a)} = \frac{(a^2+ab+b^2)(b-a)}{3(b-a)} = \frac{a^2+ab+b^2}{3}. \end{aligned}$$

4. Συνδυάζοντας τα άνω σκέλη,

$$\begin{aligned} \mathrm{VAR}(X) &= E(X^2) - (E(X))^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 \\ &= \frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12} = \frac{(b-a)^2}{12}. \end{aligned}$$

□

Παράδειγμα 8.1. Έστω Τ.Μ. X ομοιόμορφα κατανεμημένη μεταξύ των $a = 10$ και $b = 30$. Θα απαντήσουμε τα ακόλουθα ερωτήματα:

1. Ποια είναι η πιθανότητα $P(X > 25)$;

2. Ποια είναι η δεσμευμένη πιθανότητα $P(X > 25|X > 20)$;

Για να απαντήσουμε το πρώτο ερώτημα, το πιο απλό είναι να παρατηρήσουμε ότι η πιθανότητα να βρίσκεται μια ομοιόμορφη Τ.Μ. σε ένα διάστημα είναι ανάλογη του μήκους του, και επομένως

$$P(X > 25) = P(25 < X < 30) = \frac{30-25}{30-10} = \frac{1}{4}.$$

Αυστηρά, μπορούμε να ολοκληρώσουμε στο διάστημα $[25, 30]$ την πυκνότητα

$$f(x) = \begin{cases} \frac{1}{20}, & x \in [10, 30], \\ 0, & x \notin [10, 30], \end{cases}$$

οπότε λαμβάνουμε την

$$P(X > 25) = \int_{25}^{30} \frac{1}{20}\,dx = \frac{1}{4}.$$

Μπορείτε να φτάσετε στο ίδιο αποτέλεσμα χρησιμοποιώντας την κατανομή;

Σχετικά με το δεύτερο ερώτημα, το πιο απλό είναι να παρατηρήσουμε πως, αφού το ενδεχόμενο $X \le 20$ έχει αποκλειστεί, η X είναι πλέον ομοιόμορφα κατανεμημένη στο $[20, 30]$, άρα η ζητούμενη πιθανότητα είναι η

$$P(X > 25) = \frac{30-25}{30-20} = \frac{1}{2}.$$

Ο άνω διαισθητικός συλλογισμός μπορεί να γίνει απολύτως αυστηρός με χρήση της έννοιας της δεσμευμένης πυκνότητας, που δεν θα δούμε εδώ.

Εναλλακτικά, με χρήση της πυκνότητας έχουμε:

$$\begin{aligned} P(X > 20) &= \int_{20}^{30} \frac{1}{20}\, dx = \frac{1}{2}, \\ P(X > 25 | X > 20) &= \frac{P(X > 25 \text{ και } X > 20}{P(X > 20)} = \frac{P(X > 25)}{P(X > 20)} = \frac{1/4}{1/2} = \frac{1}{2}. \end{aligned}$$

Μπορείτε να φτάσετε στο ίδιο αποτέλεσμα χρησιμοποιώντας την κατανομή;

8.2 Εκθετική Κατανομή

Ορισμός 8.2. *Μια συνεχής Τ.Μ. X έχει εκθετική κατανομή με παράμετρο $\theta > 0$ αν έχει πυκνότητα*

$$f(x) = \begin{cases} \frac{1}{\theta}e^{-x/\theta}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Για συντομία, γράφουμε $X \sim$ Εκϑ(θ).

Παρατήρηση: Παρατηρήστε πως η $f(x)$ είναι παντού μη αρνητική και επιπλέον

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^{\infty} \frac{1}{\theta}e^{-x/\theta}\,dx = \int_0^{\infty} \left(-e^{-x/\theta}\right)'\,dx = \left[e^{-x/\theta}\right]_{\infty}^{0} = 1 - 0 = 1,$$

όπως πρέπει.

Λήμμα 8.2. (Ιδιότητες της εκθετικής κατανομής) *Έστω $X \sim$ Εκϑ(θ). Η X έχει τις εξής ιδιότητες:*

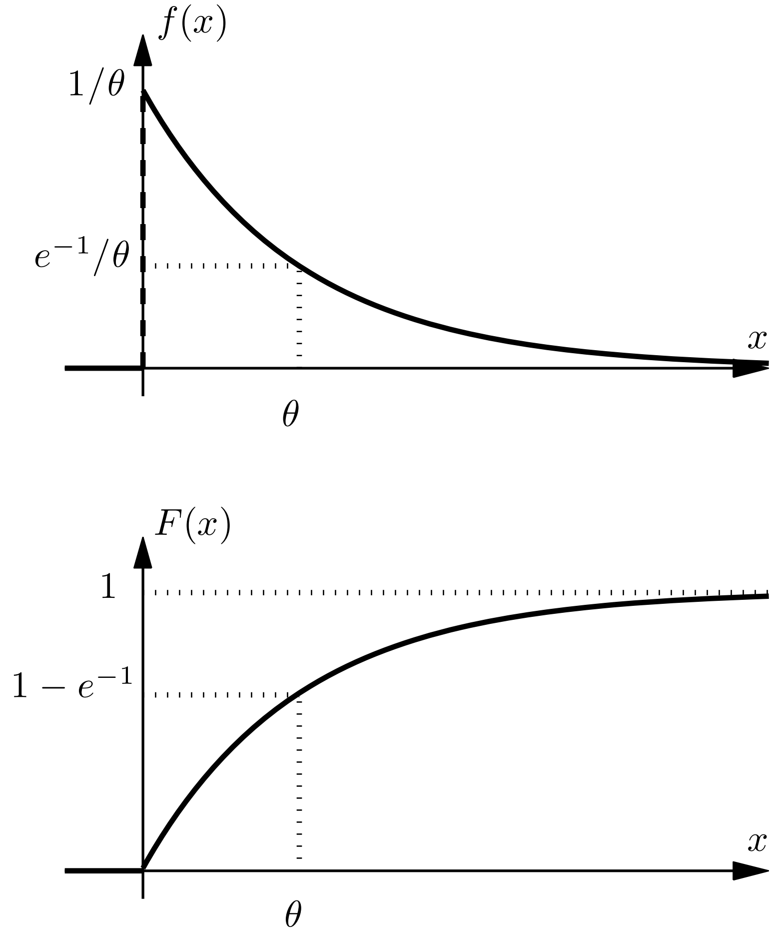
1. $F(x) = \begin{cases} 0, & x \leq 0, \\ 1 - e^{-x/\theta}, & x \geq 0. \end{cases}$
2. $E(X) = \theta$.
3. $E(X^2) = 2\theta^2$.
4. $\text{VAR}(X) = \theta^2$.
5. (Ιδιότητα έλλειψης μνήμης) *Για κάθε $a, b \geq 0$:*

$$P(X \geq a + b \,|\, X \geq a) = P(X \geq b).$$

Δηλαδή η πιθανότητα $P(X \geq a + b \,|\, X \geq a)$ δεν επηρεάζεται από το a, και είναι ίση με εκείνη που αντιστοιχεί στο $a = 0$, δηλαδή την $P(X \geq 0 + b | X \geq 0) = P(X \geq b)$.

Απόδειξη. 1. Προφανώς για $x < 0$ έχουμε $F(x) = P(X \leq x) = 0$. Για $x \geq 0$, βρίσκουμε

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)\,dy = \int_0^x \frac{1}{\theta}\,e^{-y/\theta}\,dy = \left[-e^{-y/\theta}\right]_0^x = 1 - e^{-x/\theta}.$$



Σχήμα 8.2: Γραφική παράσταση της πυκνότητας $f(x)$ και της κατανομής $F(x)$ μιας εκθετικά κατανεμημένης Τ.Μ. X με παράμετρο θ.

2. Από τον ορισμό της μέσης τιμής μιας συνεχούς Τ.Μ. και την πυκνότητα της εκθετικής κατανομής βρίσκουμε

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_0^{\infty} \frac{x}{\theta}\, e^{-x/\theta}\, dx = \int_0^{\infty} y\, e^{-y}\, \theta dy \\ &= \big[-\theta y\, e^{-y}\big]_0^{\infty} + \theta \int_0^{\infty} e^{-y}\, dy = 0 \;+\; \big[-\theta\, e^{-y}\big]_0^{\infty} = \theta, \end{aligned}$$

όπου στην τρίτη ισότητα κάναμε την αντικατάσταση $y = x/\theta$, και στην τέταρτη ισότητα ολοκληρώσαμε κατά παράγοντες, θέτοντας $u = y$ και $dv = e^{-y}dy$, έτσι ώστε $du = dy$ και $v = -e^{-y}$. Στην πέμπτη ισότητα, χρησιμοποιήσαμε το όριο

$$\big[-y\, e^{-y}\big]_0^{\infty} = -\lim_{t\to\infty} \big[y\, e^{-y}\big]_0^{t} = -\lim_{t\to\infty} \big[te^{-t} - 0\big] = -\lim_{t\to\infty} \frac{t}{e^t} = 0.$$

(Η τελευταία ισότητα είναι μια απλή εφαρμογή του κανόνα L'Hôspital.)

3. Παρομοίως,

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2\, f(x)\, dx = \int_0^{\infty} \frac{x^2}{\theta}\, e^{-x/\theta}\, dx = \int_0^{\infty} \theta y^2\, e^{-y}\, \theta dy \\ &= \big[-\theta^2 y^2\, e^{-y}\big]_0^{\infty} + 2\theta^2 \int_0^{\infty} y\, e^{-y}\, dy, \end{aligned}$$

όπου στην τρίτη ισότητα κάναμε πάλι την αντικατάσταση $y = x/\theta$, και στην τέταρτη ισότητα ολοκληρώσαμε κατά παράγοντες, θέτοντας $u = y^2$ και $dv = e^{-y}dy$, έτσι ώστε $du = 2ydy$ και $v = -e^{-y}$. Ο πρώτος όρος στην τελευταία πιο πάνω έκφραση είναι μηδενικός, όπως προκύπτει αν υπολογίσουμε το όριο με χρήση του κανόνα L'Hôspital, όπως και πριν. Επιπλέον, το τελευταίο ολοκλήρωμα ισούται με την μέση τιμή μιας Τ.Μ. με κατανομή Εκθ(1), άρα, από την πρώτη ιδιότητα, είναι ίσο με 1. Συνεπώς έχουμε

$$E(X^2) = 0 + 2\theta^2 \times 1 = 2\theta^2.$$

4. Η διασπορά της X υπολογίζεται εύκολα ως

$$\mathrm{VAR}(X) = E(X^2) - (E(X))^2 = 2\theta^2 - \theta^2 = \theta^2.$$

5. Χρησιμοποιούμε τον ορισμό της δεσμευμένης πιθανότητας και το προηγούμενο σκέλος:

$$\begin{aligned} &P(X \geq a + b \,|\, X \geq a) \\ &= \frac{P(X \geq a + b \text{ και } X \geq a)}{P(X \geq a)} = \frac{P(X \geq a + b)}{P(X \geq a)} = \frac{1 - P(X < a + b)}{1 - P(X < a)} \\ &= \frac{1 - F(a + b)}{1 - F(a)} = \frac{e^{-(a+b)/\theta}}{e^{-a/\theta}} = e^{-b/\theta} = 1 - F(b) = P(X \geq b). \end{aligned}$$

□

Παρατηρήσεις

1. Η ιδιότητα έλλειψης μνήμης εμφανίζεται και με άλλες μορφές. Για παράδειγμα,

$$\begin{aligned} P(X \leq a + b | X > a) &= P(X \leq b), \\ P(X < a + b | X \geq a) &= P(X < b). \end{aligned}$$

Σε όλες τις περιπτώσεις, η απόδειξη είναι ανάλογη της απόδειξης του λήμματος.

2. Η διαισθητική ερμηνεία της ιδιότητας έλλειψης μνήμης είναι ίδια με αυτή της περίπτωσης της γεωμετρικής κατανομής: αν δούμε την X σαν διάρκεια ζωής, τότε αν μάθουμε πως η X ξεπέρασε κάποιο όριο, η υπόλοιπη διάρκεια που απομένει έχει κατανομή ίδια με την κατανομή της X στην αρχή του χρόνου. Κατά μια έννοια, κάτι που έχει διάρκεια ζωής εκθετικά κατανεμημένη, δεν «γερνά». Παράδειγμα είναι οι πυρήνες των ραδιενεργών υλικών.

3. Από κάποιες απόψεις, η εκθετική κατανομή είναι η συνεχής περίπτωση της γεωμετρικής κατανομής. Πράγματι, μοιράζονται αρκετές από τις θεμελιώδεις ιδιότητές τους, για παράδειγμα την έλλειψη μνήμης. Το αποτέλεσμα είναι ότι έχουν παρόμοιες χρήσεις, για παράδειγμα στην μοντελοποίηση του χρόνου μέχρι να συμβεί κάτι.

   Ένας φορμαλιστικός τρόπος για να διαπιστώσουμε την ομοιότητά τους, είναι να παρατηρήσουμε πως η πυκνότητα $f(x)$ μιας Τ.Μ. με Εκθ(θ) κατανομή είναι μαθηματικά πανομοιότυπη με την μάζα $p(x)$ μιας Τ.Μ. με Γεωμ(p) κατανομή. Συγκεκριμένα, στη συνεχή περίπτωση, για $x > 0$ η πυκνότητα $f(x)$ είναι της μορφής

$$f(x) = \frac{1}{\theta} e^{-x/\theta} = C\,\gamma^x, \quad x > 0,$$

   όπου ορίσαμε τις σταθερές $C = 1/\theta$ και $\gamma = e^{-1/\theta}$. Αντίστοιχα, στη διακριτή περίπτωση, για $x \geq 1$ η μάζα $p(x)$ μπορεί να εκφραστεί ως

$$p(x) = p(1-p)^{x-1} = \frac{p}{1-p} \times (1-p)^x = C\,\gamma^x, \quad x \geq 1,$$

   όπου τώρα έχουμε τις σταθερές $C = p/(1-p)$ και $\gamma = (1-p)$. Η ομοιότητα μεταξύ των δύο κατανομών είναι προφανής.

Παράδειγμα 8.2. Έστω πως ο χρόνος X, σε μήνες, μέχρι την πρώτη φορά που ένας σκληρός δίσκος θα παρουσιάσει κάποιο σφάλμα, έχει εκθετική κατανομή με μέσο όρο 30 μήνες. Συνεπώς, η $X \sim$ Εκθ(30).

Η πιθανότητα το πρώτο σφάλμα να εμφανιστεί μετά τους πρώτους 30 μήνες είναι

$$P(X > 30) = 1 - P(X \leq 30) = 1 - F(30) = 1 - [1 - e^{-30/30}] = e^{-1} \simeq 0.3679.$$

Η πιθανότητα ο δίσκος να μην παρουσιάσει σφάλμα για τους επόμενους 30 μήνες, δεδομένου ότι ήδη λειτουργεί 30 μήνες χωρίς πρόβλημα, είναι και πάλι

$$P(X > 60 | X > 30) = P(X > 30) = e^{-1} \simeq 0.3679,$$

όπου χρησιμοποιήσαμε την ιδιότητα έλλειψης μνήμης. Τέλος, η πιθανότητα το πρώτο σφάλμα να εμφανιστεί μετά τον δέκατο μήνα αλλά πριν τον εικοστό είναι

$$P(10 < X < 20) \;=\; F(20) - F(10) = \left[1 - e^{-20/30}\right] - \left[1 - e^{-10/30}\right] \simeq 0.2031.$$

Παράδειγμα 8.3. Έστω πως εισερχόμαστε σε μια τράπεζα με 10 ταμίες και 30 πελάτες, εκ των οποίων οι 10 ήδη συναλλάσσονται με κάποιο ταμία. Οι υπόλοιποι 20 πελάτες βρίσκονται σε αυστηρή σειρά προτεραιότητας. Ποια είναι η πιθανότητα να καταφέρουμε να εξυπηρετηθούμε πριν φύγουν όλοι όσοι είναι ήδη μέσα στην τράπεζα; Υποθέτουμε ότι στην τράπεζα δεν μπαίνει ποτέ πάνω από ένας πελάτης κάθε χρονική στιγμή.

Εξετάζουμε δύο περιπτώσεις, που κατά μια έννοια είναι ακραίες. Η πρώτη είναι κάθε πελάτης να θέλει ακριβώς T χρόνο για να εξυπηρετηθεί. Σε αυτή την περίπτωση, είναι προφανές πως όταν βγούμε από την τράπεζα έχουν ήδη βγει όλοι όσοι είχαν φτάσει πριν από εμάς, γιατί έχουν ξεκινήσει την εξυπηρέτησή τους πρώτοι.

Έστω τώρα πως ο χρόνος εξυπηρέτησης είναι εκθετικός, με την ίδια παράμετρο θ για όλους τους πελάτες. Εκ πρώτης όψεως, το ερώτημα είναι πολύ δύσκολο να απαντηθεί. Η ιδιότητα όμως της έλλειψης μνήμης μας επιτρέπει να δώσουμε μια εύκολη απάντηση. Ας μεταφερθούμε χρονικά στη στιγμή που αρχίζει η δική μας εξυπηρέτηση. Λόγω της απώλειας μνήμης, και στους άλλους πελάτες που αυτή τη στιγμή εξυπηρετούνται απομένει χρόνος εξυπηρέτησης όμοια κατανεμημένος με το δικό μας, δηλαδή με εκθετική κατανομή παραμέτρου θ. Λόγω συμμετρίας, λοιπόν, η πιθανότητα να είμαστε εμείς αυτοί που θα βγούμε τελευταίοι, είναι $1/10$, άρα η ζητούμενη πιθανότητα είναι $9/10$.

Λόγω της ιδιότητας της έλλειψης μνήμης, και για ορισμένους άλλους λόγους, η εκθετική κατανομή είναι εξαιρετικά εύχρηστη σε προβλήματα της Θεωρίας Αναμονής. Αυτό το παράδειγμα είναι χαρακτηριστικό.

8.3 Κανονική Κατανομή

Ορισμός 8.3. *Μια συνεχής Τ.Μ. X έχει* <u>*κανονική (ή Γκαουσιανή) κατανομή*</u> *με παραμέτρους $\mu \in \mathbb{R}$ και $\sigma > 0$, αν έχει πυκνότητα*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \qquad x \in \mathbb{R}. \tag{8.1}$$

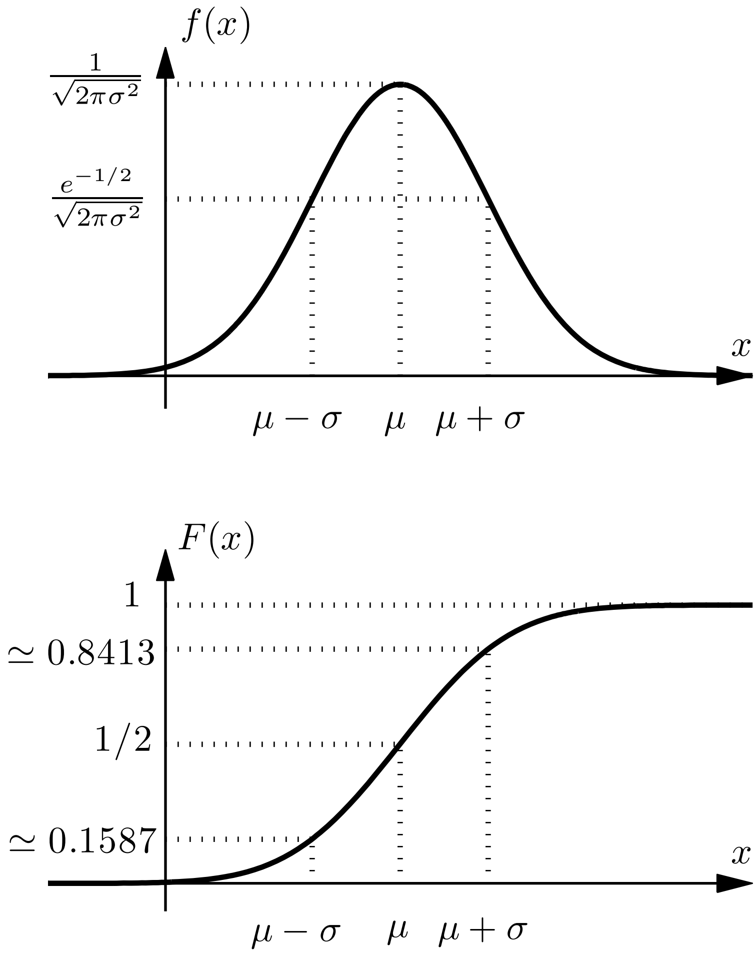
Για συντομία, γράφουμε $X \sim N(\mu, \sigma^2)$.

Παρατηρήσεις

1. Αν και, εκ πρώτης όψεως, η πολύπλοκη μορφή της πυκνότητας (8.1) ίσως μας ξενίζει, είναι σημαντικό να σημειώσουμε πως η κανονική κατανομή αντλεί τη σημασία της από το Κεντρικό Οριακό Θεώρημα, το οποίο μας λέει πως αυτή η κατανομή είναι, κατά κάποιον τρόπο, αναπόφευκτη! Συγκεκριμένα, όταν εξετάζουμε τον εμπειρικό μέσο όρο $\bar{X}_N$ ενός σχετικά μεγάλου πλήθους Τ.Μ. X_i, η κατανομή του $\bar{X}_N$ τείνει *πάντοτε*, με μια κατάλληλη κανονικοποίηση, στην κανονική κατανομή, ανεξάρτητα του τι κατανομή έχουν τα μεμονωμένα X_i!

2. Για να βεβαιωθούμε πως ο πιο πάνω ορισμός είναι μαθηματικά ορθός θα πρέπει να ελέγξουμε ότι
$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}\,dx = 1.$$
Αυτό φυσικά πράγματι ισχύει, αλλά ο σχετικός υπολογισμός είναι αρκετά τεχνικός και μακροσκελής, και για αυτό τον παραλείπουμε.

3. Έστω πως, για μια Τ.Μ. X με κατανομή $X \sim N(\mu, \sigma^2)$, θέλουμε να υπολογίσουμε την πιθανότητα $P(a \le X \le b)$. Αν και η πιθανότητα αυτή μπορεί να εκφραστεί ως το ολοκλήρωμα της πυκνότητας
$$P(a \le X \le b) = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}\,dx,$$
αυτό το ολοκλήρωμα *δεν μπορεί να εκφραστεί σε κλειστή μορφή*. Επίσης, ούτε και η κατανομή της X μπορεί να υπολογιστεί σε απλή μορφή. Σύντομα θα αναπτύξουμε μια μεθοδολογία που θα μας επιτρέπει να υπολογίζουμε τα σχετικά ολοκληρώματα με σχετικά λίγο κόπο.



Σχήμα 8.3: Γραφική παράσταση της πυκνότητας $f(x)$ και της κατανομής $F(x)$ μιας Τ.Μ. $X \sim N(\mu, \sigma^2)$.

Λήμμα 8.3. (Ιδιότητες της Κανονικής Κατανομής) *Έστω* $X \sim N(\mu, \sigma^2)$.

1. $E(X) = \mu$.
2. $\mathrm{VAR}(X) = \sigma^2$.
3. (Κανονικοποίηση): *Η Τ.Μ.* $Z = \frac{X-\mu}{\sigma}$ *έχει κατανομή* $Z \sim N(0,1)$.

Απόδειξη. 1. Παρατηρώντας πως η πυκνότητα $f(x)$ της X έχει παράγωγο

$$f'(x) = \frac{d}{dx}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} = -\frac{(x-\mu)}{\sigma^2} f(x).$$

βρίσκουμε

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_{-\infty}^{\infty} (x-\mu+\mu)\, f(x)\, dx \\
&= \sigma^2 \int_{-\infty}^{\infty} \frac{(x-\mu)}{\sigma^2}\, f(x)\, dx + \mu \int_{-\infty}^{\infty} f(x)\, dx \\
&= \sigma^2 \int_{-\infty}^{\infty} (-f'(x))\, dx + \mu = -\sigma^2 \big[f(x)\big]_{-\infty}^{\infty} + \mu = \mu.
\end{aligned}$$

2. Έχουμε:

$$\begin{aligned}
\mathrm{VAR}(X) &= E\big[(X-\mu)^2\big] = \int_{-\infty}^{\infty} (x-\mu)^2\, f(x)\, dx \\
&= \sigma^2 \int_{-\infty}^{\infty} (x-\mu)\frac{(x-\mu)}{\sigma^2}\, f(x)\, dx = \sigma^2 \int_{-\infty}^{\infty} (x-\mu)(-f'(x))\, dx \\
&= \big[-\sigma^2 (x-\mu) f(x)\big]_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} f(x)\, dx = 0 + \sigma^2 \times 1 = \sigma^2,
\end{aligned}$$

όπου στην πέμπτη ισότητα ολοκληρώσαμε κατά παράγοντες, θέτοντας $u = x - \mu$ και $dv = f'(x)dx$, έτσι ώστε $du = dx$ και $v = f(x)$. Στην έκτη ισότητα,

$$\begin{aligned}
&\big[(x-\mu) f(x)\big]_{-\infty}^{\infty} \\
&= \lim_{x\to\infty} \frac{x-\mu}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] - \lim_{x\to-\infty} \frac{x-\mu}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \\
&= \frac{1}{\sqrt{2\pi}\sigma} \left[\lim_{x\to\infty} \frac{(x-\mu)}{\exp\left[\frac{(x-\mu)^2}{2\sigma^2}\right]} - \lim_{x\to-\infty} \frac{(x-\mu)}{\exp\left[\frac{(x-\mu)^2}{2\sigma^2}\right]} \right] = 0 - 0 = 0,
\end{aligned}$$

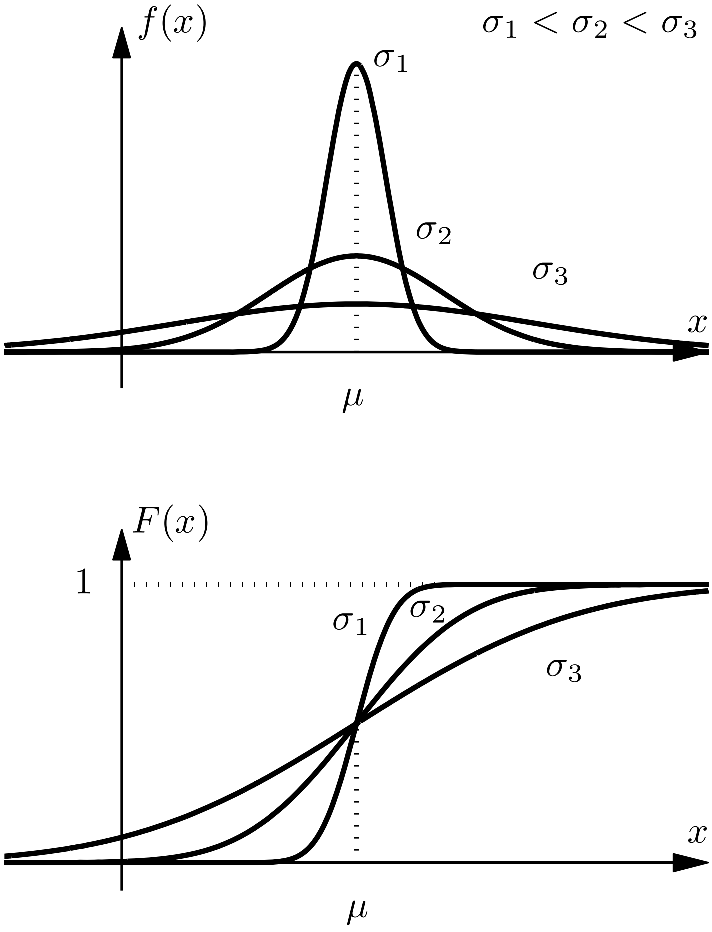
όπως προκύπτει με χρήση του κανόνα L'Hôspital.

3. Θα υπολογίσουμε την κατανομή της Z:

$$\begin{aligned}
F_Z(z) &= P(Z \le z) = P\left(\frac{X-\mu}{\sigma} \le z\right) = P(X - \mu \le \sigma z) \\
&= P(X \le \sigma z + \mu) = \int_{-\infty}^{\sigma z+\mu} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \\
&= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{t^2}{2}\right] dt.
\end{aligned}$$

Στην τελευταία ισότητα θέσαμε $t = \frac{x-\mu}{\sigma}$, άρα $x = z\sigma + \mu \Rightarrow t = z$, $x = -\infty \Rightarrow t = -\infty$, και $dx = \sigma dt$. Παρατηρήστε τώρα πως η κατανομή που βρήκαμε πράγματι ταυτίζεται με την κατανομή μιας Γκαουσιανής με παραμέτρους $\mu = 0$ και $\sigma^2 = 1$.

□

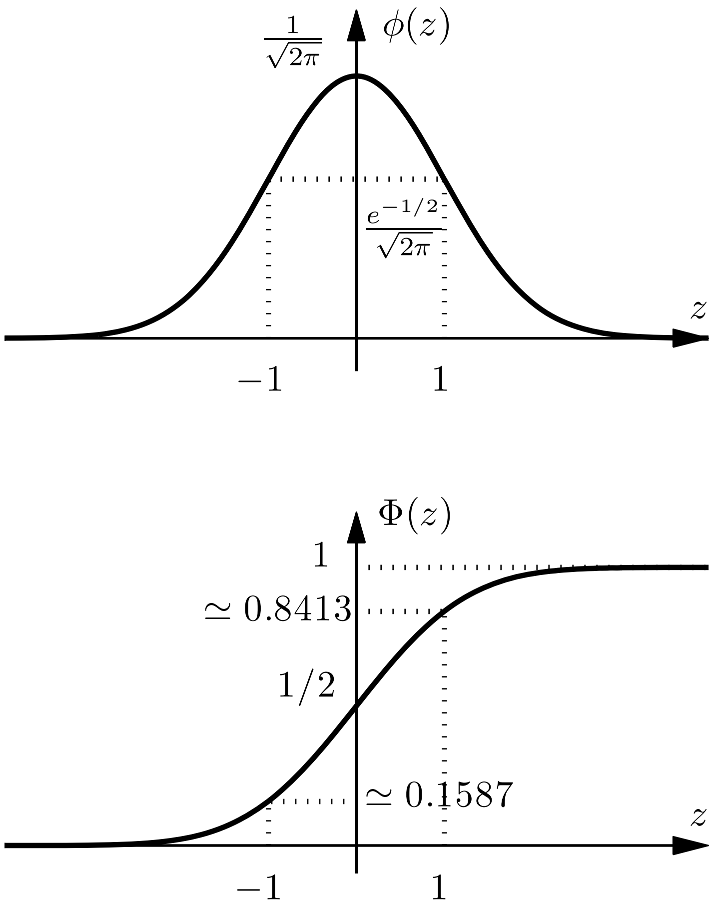


Σχήμα 8.4: Γραφική παράσταση της πυκνότητας $f(x)$ και της κατανομής $F(x)$ μιας κανονικής Τ.Μ. X για τρεις τιμές $\sigma_1 < \sigma_2 < \sigma_3$ της διασποράς.

Παρατηρήσεις

1. Στο Σχήμα 8.4 έχουμε σχεδιάσει την πυκνότητα και την κατανομή τριών κανονικών Τ.Μ. με ίδιες μέσες τιμές, αλλά αυξανόμενες διασπορές. Παρατηρήστε ότι όσο μεγαλώνει το σ^2, τόσο φαρδαίνει και χαμηλώνει η αντίστοιχη πυκνότητα, έτσι όμως ώστε το ολοκλήρωμα να παραμένει μονάδα. Ταυτοχρόνως, η κατανομή γίνεται λιγότερο απότομη.

2. Η τελευταία από τις άνω ιδιότητες μας επιτρέπει να χρησιμοποιήσουμε μια μεθοδολογία για τον υπολογισμό πιθανοτήτων που περιέχουν την Γκαουσιανή κατανομή η οποία βασίζεται στην ειδική περίπτωση που $\mu = 0, \sigma^2 = 1$. Η ειδική αυτή περίπτωση παρουσιάζει λοιπόν ξεχωριστό ενδιαφέρον, και αυτό μας οδηγεί στον ακόλουθο ορισμό.



Σχήμα 8.5: Γραφική παράσταση της πυκνότητας $\phi(z)$ και της κατανομής $\Phi(z)$ της τυπικής Τ.Μ. $N(0,1)$.

Ορισμός 8.4. *Αν η Τ.Μ. $X \sim N(0,1)$, λέμε πως έχει τυπική κανονική κατανομή και συμβολίζουμε την πυκνότητα και τη συνάρτηση κατανομής της ως*

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \qquad z \in \mathbb{R}, \tag{8.2}$$

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx, \qquad z \in \mathbb{R}, \tag{8.3}$$

αντίστοιχα.

Λήμμα 8.4. (Συμμετρία τυπικής κανονικής κατανομής). *Ισχύει το ακόλουθο:*

$$\Phi(z) = 1 - \Phi(-z).$$

Απόδειξη. Παρατηρήστε πως

$$\begin{aligned}
1-\Phi(-z) &= 1-\int_{-\infty}^{-z}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx\\
&= \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx+\int_{z}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx+\int_{\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx\\
&= \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx+\int_{z}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx-\int_{z}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx\\
&= \int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx=\Phi(z).
\end{aligned}$$

Στην τρίτη ισότητα, κάναμε αλλαγή μεταβλητής $x \to -x$ στο τρίτο ολοκλήρωμα.

□

8.4 Υπολογισμός Πιθανοτήτων Κανονικής Κατανομής

Λήμμα 8.5. (Υπολογισμός πιθανοτήτων κανονικών Τ.Μ.) *Αν* $X \sim N(\mu, \sigma^2)$, *τότε*

$$\begin{aligned} P(a \leq X \leq b) &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right), \\ P(X \leq b) &= \Phi\left(\frac{b-\mu}{\sigma}\right), \\ P(X \geq a) &= 1 - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

Απόδειξη. Παρατηρήστε πως

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

Στη δεύτερη ισότητα θέσαμε $Z = \frac{X-\mu}{\sigma}$. Η τρίτη ισότητα προκύπτει γιατί, από το τελευταίο σκέλος του Λήμματος 8.3, έχουμε $Z \sim N(0,1)$. Προκύπτει έτσι η πρώτη εξίσωση. Οι άλλες δύο προκύπτουν με ανάλογο τρόπο. □

Παράδειγμα 8.4. Αν θέλουμε να υπολογίσουμε την $P(0 \leq X \leq 1.5)$ για την $X \sim N(1,3)$, έχουμε

$$P(0 \leq X \leq 1.5) = \Phi\left(\frac{1.5-1}{\sqrt{3}}\right) - \Phi\left(\frac{0-1}{\sqrt{3}}\right) \simeq \Phi(0.2887) - \Phi(-0.5774).$$

Παρατηρήσεις

1. Αν και μπορούμε να εφαρμόσουμε πάντα το άνω λήμμα, είναι πιο εύκολο να θυμόμαστε την ακόλουθη μέθοδο, η οποία ουσιαστικά ακολουθεί τα βήματα της απόδειξης του λήμματος.

   Έστω πως θέλουμε να υπολογίσουμε κάποια πιθανότητα της μορφής $P(a \leq X \leq b)$ για μια Τ.Μ. $X \sim N(\mu, \sigma^2)$. (Η μέθοδος τροποποιείται με τον προφανή τρόπο προκειμένου να υπολογίσουμε πιθανότητες της μορφής $P(X \leq b)$ ή $P(X \geq a)$.)

   Καταρχήν, εκφράζουμε την ζητούμενη πιθανότητα ως

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right).$$

Π.χ., αν η ζητούμενη πιθανότητα ήταν η $P(0 \leq X \leq 1.5)$ για την $X \sim N(1,3)$, όπως στο άνω παράδειγμα, τότε

$$\begin{aligned} & P(0 \leq X \leq 1.5) \\ = \; & P\left(\frac{0-\mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{1.5-\mu}{\sigma}\right) = P\left(\frac{0-1}{\sqrt{3}} \leq \frac{X-\mu}{\sigma} \leq \frac{1.5-1}{\sqrt{3}}\right) \\ = \; & P\left(-\frac{1}{\sqrt{3}} \leq \frac{X-\mu}{\sigma} \leq \frac{0.5}{\sqrt{3}}\right). \end{aligned}$$

Από την Ιδιότητα 3 του Λήμματος 8.3 προκύπτει πως η $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, άρα έχουμε

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

Στο άνω παράδειγμα,

$$\begin{aligned} P(0 \leq X \leq 1.5) &= P\left(-\frac{1}{\sqrt{3}} \leq Z \leq \frac{0.5}{\sqrt{3}}\right) = \Phi\left(\frac{0.5}{\sqrt{3}}\right) - \Phi\left(-\frac{1}{\sqrt{3}}\right) \\ &\simeq \Phi(0.2887) - \Phi(-0.5774). \end{aligned}$$

2. Ανεξάρτητα από το αν χρησιμοποιήσουμε το άνω λήμμα ή την άνω μέθοδο, τελικά χρειαζόμαστε τις τιμές της Φ για ένα ή περισσότερα ορίσματα. Τις τιμές αυτές μπορούμε να τις βρούμε με χρήση του Πίνακα 8.1 (ή κάποιου αντίστοιχου — όλα τα βιβλία πιθανοτήτων έχουν ένα τέτοιο πίνακα). Ο Πίνακας 8.1 περιέχει, για ένα μεγάλο εύρος θετικών ορισμάτων z, τις αντίστοιχες τιμές $\Phi(z)$, ως εξής: Οι γραμμές δίνουν το ακέραιο μέρος και το πρώτο δεκαδικό ψηφίο του z και οι στήλες το δεύτερο δεκαδικό του ψηφίο. Το αντίστοιχο στοιχείο του πίνακα δίνει το $\Phi(z)$. Καθώς $\Phi(z) = 1 - \Phi(-z)$, μπορούμε να προσδιορίσουμε την τιμή της Φ και για αρνητικά ορίσματα.

   Στην περίπτωση του άνω παραδείγματος, από την τρίτη γραμμή και την τελευταία στήλη βρίσκουμε $\Phi(0.2887) \simeq \Phi(0.29) \simeq 0.6141$. Επιπλέον, για να βρούμε την τιμή $\Phi(-0.5774) \simeq \Phi(-0.58) = 1 - \Phi(0.58)$, πάμε στον πίνακα και κοιτάμε στην έκτη γραμμή και στην ένατη στήλη και βρίσκουμε πως $\Phi(0.58) \simeq 0.7190 \Rightarrow \Phi(-0.58) \simeq 0.2810$. Τελικά έχουμε

$$P(0 \leq X \leq 1.5) \simeq \Phi(0.29) - \Phi(-0.58) \simeq 0.6164 - 0.2810 = 0.3331.$$

Πίνακας 8.1: Τιμές της τυπικής κανονικής συνάρτησης κατανομής $\Phi(z)$ για θετικά ορίσματα z. Οι γραμμές δίνουν το ακέραιο μέρος και το πρώτο δεκαδικό ψηφίο, και οι στήλες το δεύτερο δεκαδικό ψηφίο του ορίσματος z. Για αρνητικά ορίσματα, χρησιμοποιούμε τη σχέση $\Phi(z) = 1 - \Phi(-z)$.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Παράδειγμα 8.5. Έστω πως το ρολόι ενός επεξεργαστή αποκλίνει κατά X δευτερόλεπτα μετά από ένα χρόνο συνεχούς λειτουργίας, όπου $X \sim N(1,4)$. Ποια η πιθανότητα μετά από ένα χρόνο:

1. Το ρολόι να πάει «πίσω», δηλαδή να έχει αρνητική απόκλιση;

2. Η απόκλιση να είναι μεγαλύτερη από $+3.5$ δευτερόλεπτα;

3. Η απόκλιση να είναι μικρότερη από ± 2 δευτερόλεπτα;

1. Χρησιμοποιώντας εδώ την γενική πιο πάνω μέθοδο με $\mu = 1$ και $\sigma = \sqrt{4} = 2$, ορίζοντας μια τυπική κανονική Τ.Μ. $Z \sim N(0,1)$, για το πρώτο ερώτημα εύκολα βρίσκουμε

$$P(X \leq 0) = P\left(\frac{X-1}{2} \leq \frac{0-1}{2}\right) = P(Z \leq -0.5) = \Phi(-0.5),$$

και αντικαθιστώντας την τιμή της $\Phi(-0.5) \simeq 0.3085$ με χρήση του Πίνακα 8.1, έχουμε $P(X \leq 0) \simeq 0.3085$. Δείτε την πρώτη από τις γραφικές παραστάσεις του Σχήματος 8.6.

2. Παρομοίως,

$$\begin{aligned} P(X > 3.5) &= P\left(\frac{X-1}{2} > \frac{3.5-1}{2}\right) = P(Z > 1.25) = 1 - P(Z \leq 1.25) \\ &= 1 - \Phi(1.25) \simeq 1 - 0.8944 = 0.1056, \end{aligned}$$

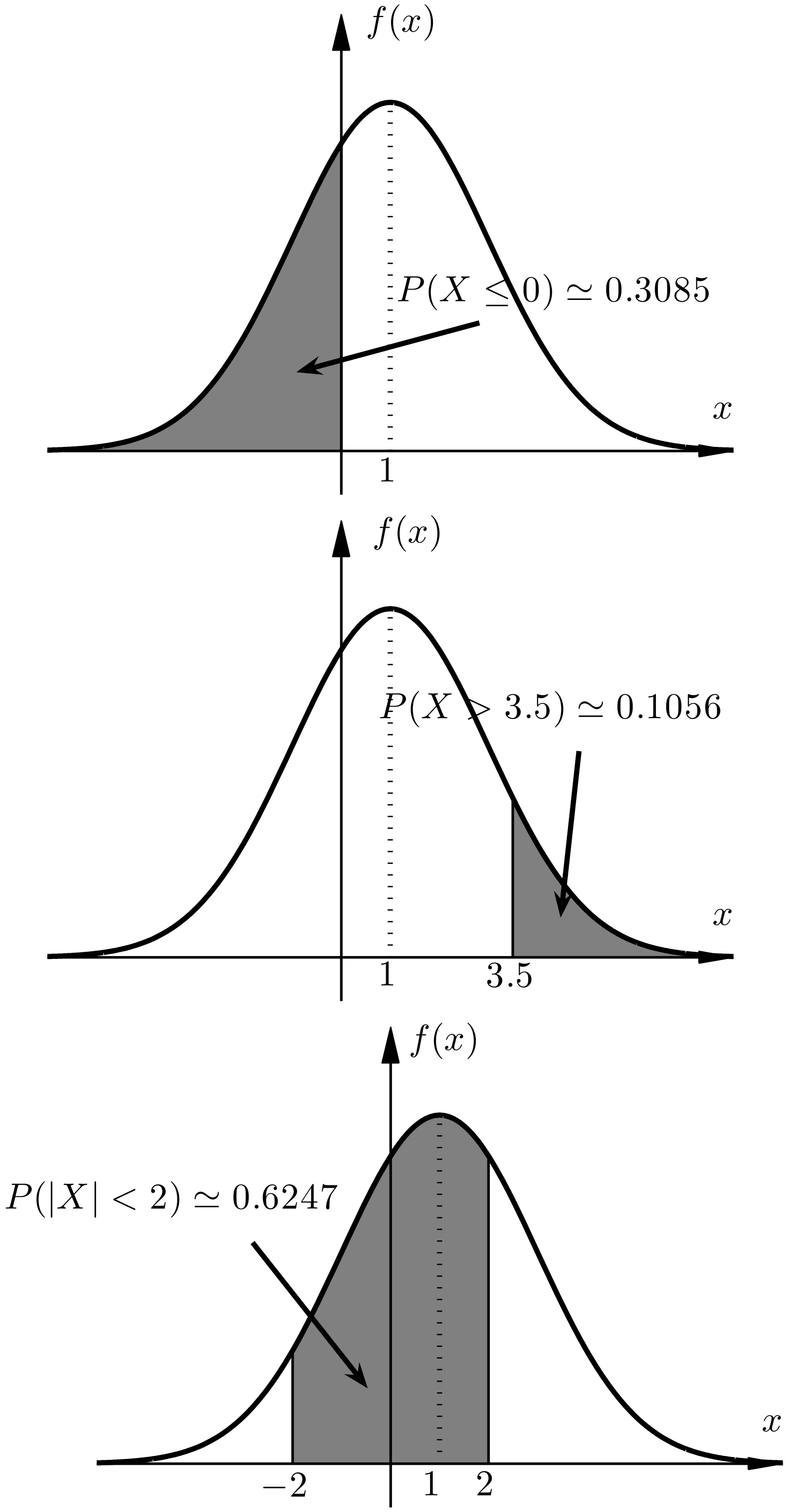
όπου αντικαταστήσαμε την τιμή της $\Phi(1.25) \simeq 0.8944$ από τον Πίνακα 8.1. Δείτε την δεύτερη από τις γραφικές παραστάσεις του Σχήματος 8.6.

3. Τέλος,

$$\begin{aligned} P(|X| < 2) &= P(-2 < X < 2) = P\left(\frac{-2-1}{2} < \frac{X-1}{2} < \frac{2-1}{2}\right) \\ &= P(-1.5 < Z < 0.5) = \Phi(0.5) - \Phi(-1.5) \\ &\simeq 0.6915 - 0.0668 = 0.6247, \end{aligned}$$

όπου αντικαταστήσαμε τις τιμές $\Phi(0.5) \simeq 0.6915$ και $\Phi(-1.5) \simeq 0.0668$ με χρήση του Πίνακα 8.1. Δείτε την τρίτη από τις γραφικές παραστάσεις του Σχήματος 8.6.

Παράδειγμα 8.6. Έστω πως η ένταση του ηλεκτρικού θορύβου σε ένα κύκλωμα είναι $X \sim N(-2,3)$. Ποια η πιθανότητα το X να μην ξεπερνάει το 1 σε απόλυτη τιμή;



Σχήμα 8.6: Οι πιθανότητες που υπολογίστηκαν στο Παράδειγμα 8.5 εμφανίζονται ως σκιασμένα εμβαδά.

Όπως και στο προηγούμενο παράδειγμα, εφαρμόζοντας την γενική πιο πάνω μέθοδο με $\mu = -2$ και $\sigma = \sqrt{3}$ και ορίζοντας μια τυπική κανονική Τ.Μ. $Z \sim N(0,1)$, βρίσκουμε

$$\begin{aligned} P(|X| \leq 1) &= P(-1 \leq X \leq 1) = P\left(\frac{-1-(-2)}{\sqrt{3}} \leq \frac{X-(-2)}{\sqrt{3}} \leq \frac{1-(-2)}{\sqrt{3}}\right) \\ &\simeq P(0.58 \leq Z \leq 1.73) = \Phi(1.73) - \Phi(0.58) \\ &\simeq 0.9582 - 0.7190 \ = \ 0.2392, \end{aligned}$$

όπου αντικαταστήσαμε τις τιμές $\Phi(1.73) \simeq 0.9582$ και $\Phi(0.58) \simeq 0.7190$ από τον Πίνακα 8.1.

Παράδειγμα 8.7. *(Τυχαία επιλογή Τ.Μ.)* Η Τ.Μ. X έχει ομοιόμορφη κατανομή στο διάστημα $[-3, 5]$, και η Τ.Μ. Y έχει κανονική κατανομή με μέση τιμή 1 και διασπορά 10. Ρίχνουμε ένα δίκαιο κέρμα, και αν έρθει κορώνα καταγράφουμε την τιμή του X ενώ αν έρθει γράμματα καταγράφουμε την τιμή του Y. Η ρίψη του κέρματος είναι ανεξάρτητη των τιμών των X, Y. Αν η τιμή που καταγράψαμε είναι μεγαλύτερη του 4, ποια είναι η πιθανότητα να είχαμε επιλέξει την X;

Για να απαντήσουμε το ερώτημα, ορίζουμε την βοηθητική Τ.Μ. Z η οποία παίρνει την τιμή 1 αν φέρουμε κορώνα και την τιμή 0 αν φέρουμε γράμματα. Επιπλέον, έστω T η τιμή που καταγράψαμε. Χρησιμοποιώντας τον κανόνα του Bayes, η ζητούμενη πιθανότητα ισούται με:

$$\begin{aligned} P(Z=1|T>4) &= \frac{P(Z=1, T>4)}{P(T>4)} \\ &= \frac{P(Z=1, X>4)}{P(T>4|Z=1)P(Z=1) + P(T>4|Z=0)P(Z=0)} \\ &= \frac{P(X>4)P(Z=1)}{P(X>4)P(Z=1) + P(Y>4)P(Z=0)} \\ &= \frac{\frac{1}{2}\int_4^5 \frac{1}{8}\,dx}{\frac{1}{2}\int_4^5 \frac{1}{8}\,dx + \frac{1}{2}P\left(\frac{Y-1}{\sqrt{10}} > \frac{4-1}{\sqrt{10}}\right)} = \frac{\frac{1}{8}}{\frac{1}{8} + \left(1 - \Phi\left(\frac{3}{\sqrt{10}}\right)\right)} \simeq 0.422. \end{aligned}$$

Στον αριθμητή της δεύτερης γραμμής χρησιμοποιήσαμε το ότι αν $Z = 1$, τότε $T = X$. Στον αριθμητή της τρίτης γραμμής χρησιμοποιήσαμε την ανεξαρτησία των X, Z. Στον παρονομαστή της τρίτης γραμμής χρησιμοποιήσαμε το γεγονός ότι αν $Z = 1$, τότε $T = X$, ενώ αν $Z = 0$, τότε $T = Y$.

Παράδειγμα 8.8. *(Ύψη μαθητών)* Το ύψος X των μαθητών ενός σχολείου ακολουθεί την Γκαουσιανή κατανομή με μέση τιμή $m_A = 1.70\ m$ για τα αγόρια και $m_K = 1.60\ m$ για τα κορίτσια, και διασπορά $\sigma^2 = (0.1\ m)^2$, κοινή για αγόρια και

κορίτσια. Η πιθανότητα ένας μαθητής να είναι κορίτσι είναι 0.5. Ποια είναι η πιθανότητα το ύψος ενός μαθητή που επιλέγεται στην τύχη να ξεπερνά το 1.80 m;

Έστω τα ενδεχόμενα A =«Ο μαθητής είναι αγόρι» και B =«ο μαθητής είναι πάνω από 1.80 m». Έχουμε:

$$\begin{aligned} P(B) &= P(B|A)P(A) + P(B|A')P(A') = \frac{1}{2}P(X > 1.8|A) + \frac{1}{2}P(X > 1.8|A') \\ &= \frac{1}{2}P\left(\frac{X-1.7}{0.1} > 1\middle|A\right) + \frac{1}{2}P\left(\frac{X-1.6}{0.1} > 2\middle|A'\right) \\ &= \frac{1}{2}(1-\Phi(1)) + \frac{1}{2}(1-\Phi(2)) \simeq 0.0907. \end{aligned}$$

Στα άνω, χρησιμοποιήσαμε το ότι αν ο μαθητής είναι αγόρι, το ύψος έχει μέση τιμή 1.7 ενώ, αν είναι κορίτσι, το ύψος έχει μέση τιμή 1.6.

8.5 Μετασχηματισμοί $Y = f(X)$

Συχνά σε εφαρμογές μας δίνεται η πυκνότητα ή η κατανομή μιας Τ.Χ. X και πρέπει να βρούμε την πυκνότητα ή την κατανομή μιας συνάρτησής της, έστω $Y = f(X)$. Δείτε τα ακόλουθα παραδείγματα.

Παράδειγμα 8.9. *(Ομοιόμορφη → εκθετική)* Έστω μια Τ.Μ. U ομοιόμορφα κατανεμημένη στο $(0, 1]$. Αν $\lambda > 0$, η νέα Τ.Μ. $Y = -\lambda \log U$ έχει κατανομή Εκθ(λ). Για να δείξουμε αυτή την ιδιότητα, θα υπολογίσουμε την συνάρτηση κατανομής $F(y)$ της Y. Για $y \geq 0$,

$$\begin{aligned} F(y) &= P(Y \leq y) = P(-\lambda \log U \leq y) = P(\log U \geq -y/\lambda) \\ &= P(U \geq e^{-y/\lambda}) = \int_{e^{-y/\lambda}}^{1} 1\, dx = 1 - e^{-y/\lambda}. \end{aligned}$$

Για $y < 0$,

$$F(y) = P(Y \leq y) = P(-\lambda \log U \leq y) = P(\log U \geq -y/\lambda) = 0.$$

Συνεπώς, πράγματι $Y \sim$ Εκθ(λ). Δείτε το Σχήμα 8.7 για μια γραφική απεικόνιση του αποτελέσματος.

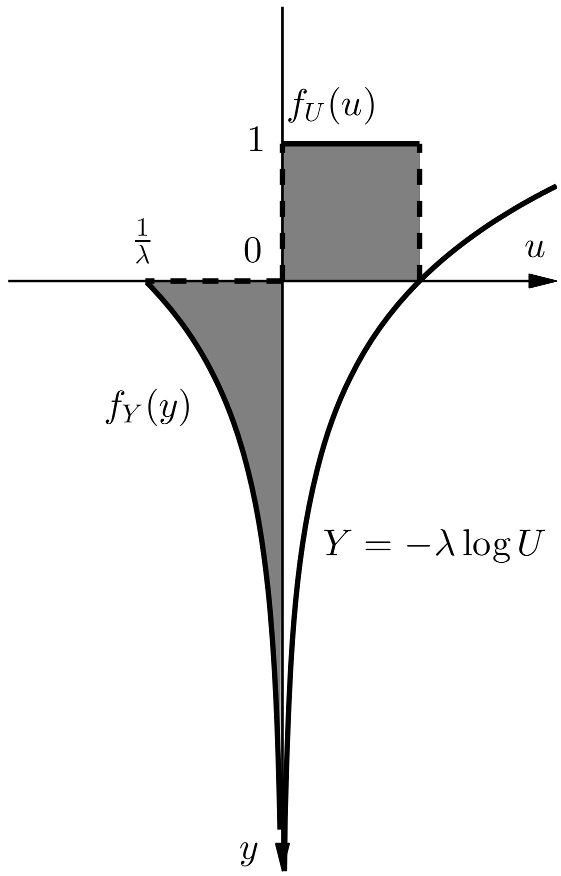
Παράδειγμα 8.10. *(Δημιουργία οποιασδήποτε κατανομής)* Έστω $F(\cdot)$ συνεχής κατανομή και έστω U Τ.Μ. ομοιόμορφα κατανεμημένη στο διάστημα $[0, 1]$. Αν $X = F^{-1}(U)$, να δειχθεί ότι η X ακολουθεί την κατανομή F. Ποια είναι η πρακτική σημασία αυτού του αποτελέσματος;

Παρατηρούμε απλά πως

$$P(X \leq a) = P(F^{-1}(U) \leq a) = P(U \leq F(a)) = \int_0^{F(a)} 1\, dx = F(a).$$

Για την δεύτερη εξίσωση χρησιμοποιήσαμε το γεγονός ότι η F είναι αύξουσα, συνεπώς και η αντίστροφή της F^{-1} είναι αύξουσα. Για την τρίτη εξίσωση χρησιμοποιήσαμε το ότι η U είναι ομοιόμορφα κατανεμημένη στο $(0, 1)$.

Το αποτέλεσμα έχει τεράστια πρακτική σημασία. Αν μπορούμε να δημιουργήσουμε με κάποιο τρόπο μια Τ.Μ. ομοιόμορφα κατανεμημένη στο $[0, 1]$, αυτόματα μπορούμε να δημιουργήσουμε οποιαδήποτε άλλη. Σχεδόν όλες οι γλώσσες προγραμματισμού έχουν μια ρουτίνα που επιστρέφει μια Τ.Μ. ομοιόμορφα κατανεμημένη στο $[0, 1]$, που συνήθως ονομάζεται rand(), αλλά ελάχιστες (πέραν των εξειδικευμένων) παρέχουν ρουτίνες που να επιστρέφουν Τ.Μ. με άλλες κατανομές. Ένας βασικός θεωρητικός λόγος για αυτή την έλλειψη είναι το αποτέλεσμα αυτής της άσκησης: αν έχουμε στη διάθεσή μας την rand(), μπορούμε εύκολα να φτιάξουμε όποια άλλη κατανομή θέλουμε!



Σχήμα 8.7: Παράδειγμα 8.9. Αν επιλέξουμε ομοιόμορφα έναν αριθμό μέσα στο διάστημα $[0, 1]$ του άξονα u, τότε επιλέγουμε επίσης έναν εκθετικά κατανεμημένο αριθμό από τον θετικό ημιάξονα y.

Παράδειγμα 8.11. *(Τετράγωνο εκθετικής)* Έστω μια Τ.Μ. X με κατανομή Εκθ(1). Έστω $f(x)$ και $F(x)$ η πυκνότητα και η κατανομή της αντίστοιχα. Έστω μια νέα Τ.Μ. $Y = aX^2$, όπου $a > 0$. Θα βρούμε την πυκνότητά της $g(y)$ και την κατανομή της $G(y)$.

Για $y < 0$ έχουμε $G(y) = P(Y \leq y) = P(aX^2 \leq y) = 0$, άρα $g(y) = 0$ για $y < 0$.

Υπολογίζουμε κατόπιν την $G(y)$ για $y \geq 0$:

$$G(y) = P(Y \leq y) = P(aX^2 \leq y) = P(X \leq \sqrt{y/a}) = F(\sqrt{y/a}) = 1 - e^{-\sqrt{y/a}}.$$

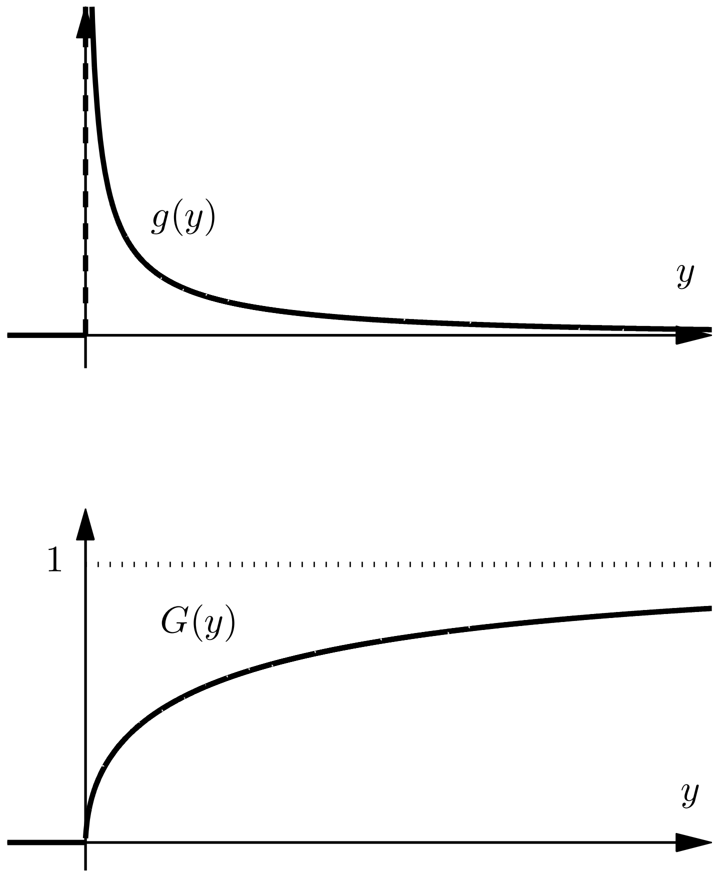
Η πυκνότητα της Y για $y > 0$ προκύπτει ως εξής:

$$g(y) = G'(y) = \left(1 - e^{-\sqrt{y/a}}\right)' = \frac{1}{2\sqrt{ay}} e^{-\sqrt{y/a}},$$

ενώ η τιμή της στο 0 δεν έχει σημασία. Συγκεντρωτικά,

$$G(y) = \begin{cases} 0, & y < 0, \\ 1 - e^{-\sqrt{y/a}}, & y \geq 0, \end{cases} \qquad g(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{2\sqrt{ay}} e^{-\sqrt{y/a}}, & y \geq 0. \end{cases}$$

Οι $g(y)$, $G(y)$ έχουν σχεδιαστεί στο Σχήμα 8.8.



Σχήμα 8.8: Γραφική παράσταση της πυκνότητας $g(y)$ και της κατανομής $G(y)$ της Τ.Μ. y στο Παράδειγμα 8.11.

Παράδειγμα 8.12. *(Χρόνος μετάδοσης)* Ο χρόνος T που μεσολαβεί από την αποστολή μέχρι την παράδοση ενός email μεγέθους X KB από έναν διακομιστή σε έναν άλλον είναι $T = X(X+1)/4$ λεπτά. Αν το X έχει εκθετική κατανομή με μέση τιμή 5 KB, θα υπολογίσουμε τη μέση τιμή του χρόνου T και την πιθανότητα ο χρόνος εκτέλεσης να ξεπεράσει τα 10 λεπτά. Έχουμε:

1. Η μέση τιμή του χρόνου T είναι:

$$\begin{aligned} E(T) &= E\left(\frac{X(X+1)}{4}\right) = \frac{1}{4}E(X^2+X) = \frac{1}{4}[E(X^2)+E(X)] \\ &= \frac{1}{4}[\mathrm{VAR}(X) + (E(X))^2 + E(X)] = \frac{1}{4}[5^2+5^2+5] = 13.75, \end{aligned}$$

όπου στην τέταρτη ισότητα χρησιμοποιήσαμε το ότι $\mathrm{VAR}(X) = E(X^2) - (E(X))^2$, και στο τέλος αντικαταστήσαμε τις γνωστές τιμές των παραμέτρων για την εκθετική κατανομή με παράμετρο 5.

2. Η πιθανότητα ο χρόνος εκτέλεσης να ξεπεράσει τα 10 λεπτά, δηλαδή $T > 10$, είναι

$$\begin{aligned}
&P(T > 10)\\
&= P\left(\frac{X(X+1)}{4} > 10\right) = P(X(X+1) > 40) = P(X^2 + X - 40 > 0)\\
&= P\left(\left(X - \frac{-1-\sqrt{161}}{2}\right)\left(X - \frac{-1+\sqrt{161}}{2}\right) > 0\right)\\
&= P\left(X > \frac{-1+\sqrt{161}}{2}\right) = 1 - P\left(X \leq \frac{-1+\sqrt{161}}{2}\right)\\
&= e^{-(-1+\sqrt{161})/10} \simeq 0.3107,
\end{aligned}$$

όπου στην πέμπτη ισότητα χρησιμοποιήσαμε το γεγονός ότι πάντοτε το X είναι θετικό, και στο τέλος αντικαταστήσαμε τον τύπο της συνάρτησης κατανομής μιας εκθετικής Τ.Μ. με παράμετρο 5.

Παρατήρηση: Σε περίπτωση που η $f(X)$ είναι γραμμική, δηλαδή $Y = aX + b$, μπορούμε να χρησιμοποιήσουμε το ακόλουθο λήμμα:

Λήμμα 8.6. (Γραμμικός μετασχηματισμός μιας Τ.Μ.) *Έστω συνεχής Τ.Μ. X με πυκνότητα $f(x)$ και κατανομή $F(x)$. Έστω $a, b \in \mathbb{R}$, με $a \neq 0$. Η Τ.Μ. $Y = aX + b$ είναι επίσης συνεχής, με πυκνότητα $g(y)$ και κατανομή $G(y)$ που ισούνται με*

$$\begin{aligned}
g(y) &= \frac{1}{|a|} f\left(\frac{y-b}{a}\right), \\
G(y) &= F(\frac{y-b}{a}), && \text{αν } a > 0,\\
G(y) &= 1 - F(\frac{y-b}{a}), && \text{αν } a < 0.
\end{aligned}$$

Απόδειξη. Θα μελετήσουμε χωριστά τις δύο περιπτώσεις: $a > 0$ και $a < 0$.

Έστω λοιπόν $a > 0$. Παρατηρούμε πως

$$G(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F\left(\frac{y-b}{a}\right).$$

Επιπλέον,

$$g(y) = G'(y) = \frac{1}{a}F'\left(\frac{y-b}{a}\right) = \frac{1}{|a|}f\left(\frac{y-b}{a}\right).$$

Έστω τώρα πως $a < 0$. Παρατηρούμε πως

$$G(y) = P(Y \le y) = P(aX + b \le y) = P\left(X \ge \frac{y-b}{a}\right) = 1 - F\left(\frac{y-b}{a}\right).$$

Επιπλέον,

$$g(y) = G'(y) = \frac{d}{dy}\left[1 - F\left(\frac{y-b}{a}\right)\right] = -\frac{1}{a}F'\left(\frac{y-b}{a}\right) = \frac{1}{|a|}f\left(\frac{y-b}{a}\right),$$

και η απόδειξη ολοκληρώθηκε. □

Παράδειγμα 8.13. Έστω $X \sim$ Εκθ(1), δηλαδή με πυκνότητα και κατανομή

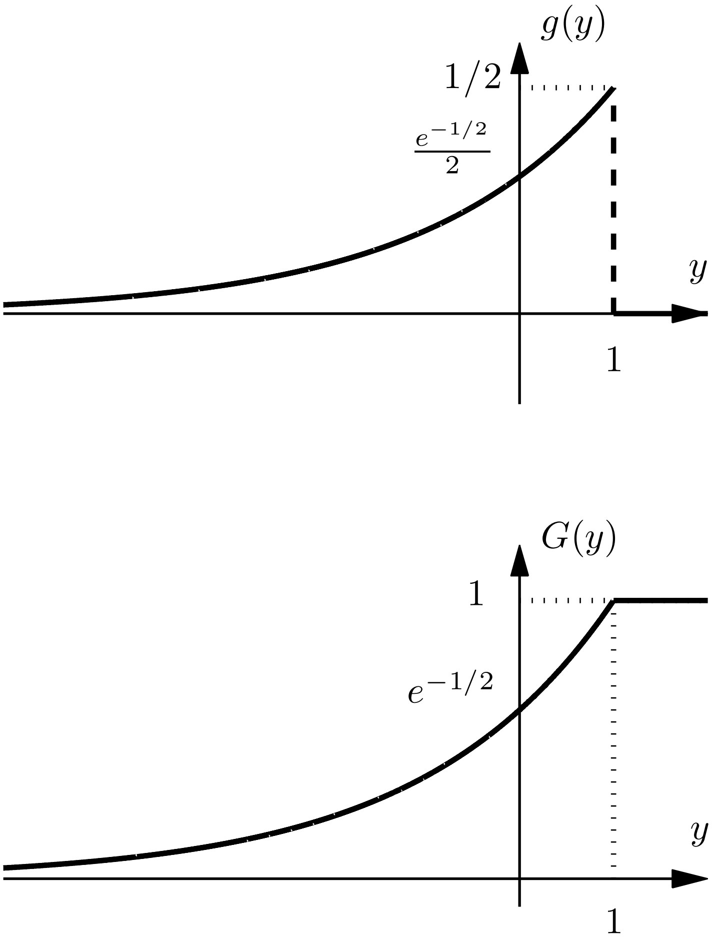
$$f(x) = \begin{cases} e^{-x}, & x \ge 0, \\ 0, & x < 0, \end{cases} \qquad F(x) = \begin{cases} 1 - e^{-x}, & x \ge 0, \\ 0, & x < 0. \end{cases}$$

Έστω επίσης η Τ.Μ. $Y = 1 - 2X$. Η πυκνότητα και η κατανομή της θα ισούνται με

$$\begin{aligned} g(y) &= \frac{1}{2}f\left(-\frac{y-1}{2}\right) = \begin{cases} \frac{1}{2}e^{\frac{y-1}{2}}, & y \le 1, \\ 0, & y > 1, \end{cases} \\ G(y) &= 1 - F\left(-\frac{y-1}{2}\right) = \begin{cases} e^{\frac{y-1}{2}}, & y \le 1, \\ 1, & y > 1. \end{cases} \end{aligned}$$

Οι $g(y), G(y)$ έχουν σχεδιαστεί στο Σχήμα 8.9.

Παρατήρηση: Πολύ συχνά όμως η f δεν είναι γραμμική συνάρτηση. Δεν υπάρχει κάποιο εύχρηστο γενικό θεώρημα για αυτή την περίπτωση, που να μας επιτρέπει τον υπολογισμό της κατανομής και της πυκνότητας της $f(X)$, ή έστω πιθανοτήτων ενδεχόμενων που αφορούν την $f(X)$. Όμως, μπορεί πάντα να εφαρμοστεί η γενική μεθοδολογία που εφαρμόστηκε στην απόδειξη του Λήμματος 8.6 και στα Παραδείγματα 8.9 έως 8.12.



Σχήμα 8.9: Γραφική παράσταση της πυκνότητας $g(y)$ και της κατανομής $G(y)$ της Τ.Μ. Y στο Παράδειγμα 8.13.

Κεφάλαιο 9

Ζεύγη Συνεχών Τυχαίων Μεταβλητών

Ανάλογα με ό,τι συμβαίνει στην περίπτωση των διακριτών Τ.Μ., η σχέση που έχουν μεταξύ τους δύο συνεχείς Τ.Μ. δεν μπορεί να αποτυπωθεί με χρήση των πυκνοτήτων τους, και είναι απαραίτητο να εισάγουμε την έννοια της *από κοινού* πυκνότητας. Μια βασική διαφορά της θεωρίας που προκύπτει σε σχέση με την περίπτωση της μελέτης μεμονωμένων συνεχών Τ.Μ., είναι ότι τα ολοκληρώματα που εμφανίζονται στον υπολογισμό πιθανοτήτων δεν είναι πια απλά, αλλά διπλά, και συνεπώς είναι αρκετά πιο σύνθετα στον υπολογισμό. Ξεκινάμε με μια σύντομη αναφορά στα διπλά ολοκληρώματα και στο βασικό θεώρημα που μας επιτρέπει τον υπολογισμό τους, το Θεώρημα του Fubini. Στο υπόλοιπο κεφάλαιο, η ανάπτυξη της θεωρίας είναι ανάλογη με αυτή του αντίστοιχου κεφαλαίου για τα ζεύγη διακριτών Τ.Μ.

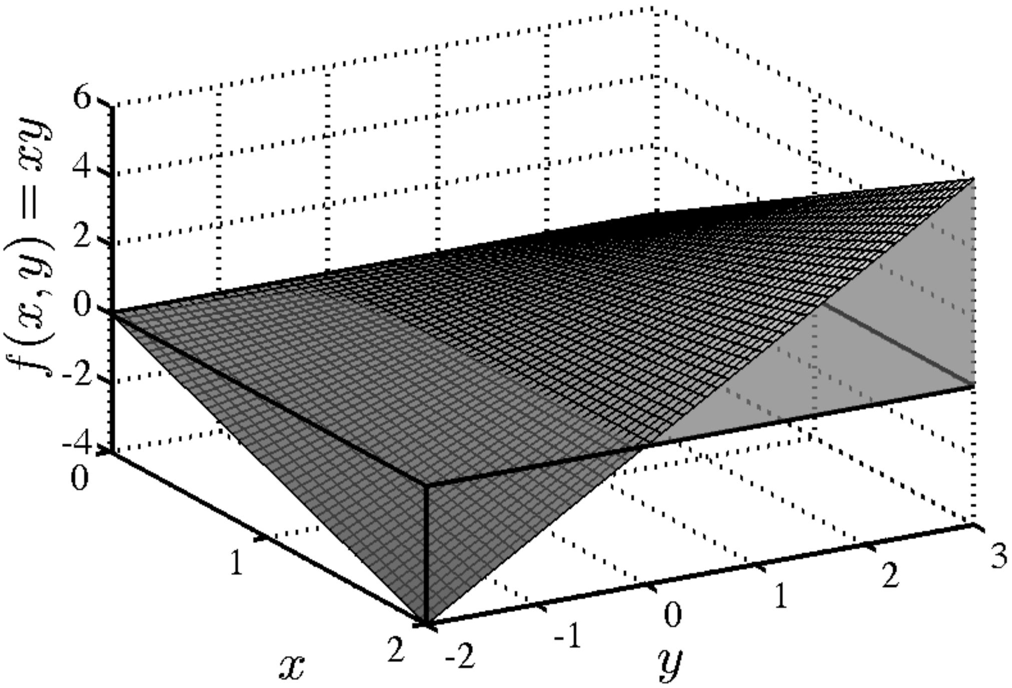
9.1 Διπλά Ολοκληρώματα

Στην παράγραφο αυτή θα αναφερθούμε, πολύ συνοπτικά, σε στοιχεία της θεωρίας των διπλών ολοκληρωμάτων που θα χρειαστούμε σε αυτό το κεφάλαιο.

Παραλείπουμε τον αυστηρό ορισμό του διπλού ολοκληρώματος, που είναι αρκετά εκτεταμένος. Πρακτικά, το διπλό ολοκλήρωμα της $f(x, y)$ στο χωρίο $R \subset \mathbb{R} \times \mathbb{R}$ συμβολίζεται με

$$\iint_R f(x,y)\, dA$$

και ισούται με τον όγκο του στερεού που περικλείεται μεταξύ του R και της $f(x)$, όπου αυτή είναι θετική, μείον τον όγκο του στερεού που περικλείεται μεταξύ του R και της $f(x)$, όπου αυτή είναι αρνητική. Δηλαδή, ενώ τα απλά ολοκληρώματα εκφράζουν προσημασμένο *εμβαδόν*, τα διπλά εκφράζουν προσημασμένο *όγκο*. Αυτή η ερμηνεία είναι αρκετή για τις δικές μας ανάγκες. Για παράδειγμα, το διπλό ολοκλήρωμα της συνάρτησης $f(x, y) = xy$ στο χωρίο $R = [0, 2] \times [-2, 3]$ είναι ο όγκος του στερεού που βρίσκεται μεταξύ του χωρίου $R_1 = [0, 2] \times [0, 3]$, όπου η συνάρτηση είναι θετική,



Σχήμα 9.1: Το διπλό ολοκλήρωμα της $f(x,y) = xy$ ισούται με τον όγκο του στερεού πάνω από το χωρίο $R = [0,2] \times [0,3]$ μείον τον όγκο του στερεού κάτω από το χωρίο $R = [0,2] \times [-2,0]$.

και του γραφήματος της $f(x,y)$, μείον τον όγκο του στερεού που βρίσκεται μεταξύ του χωρίου $R_2 = [0,2] \times [-2,0]$, όπου η συνάρτηση είναι αρνητική, και του γραφήματος της $f(x,y)$. Δείτε το Σχήμα 9.1. Συχνά, βέβαια, το χωρίο R έχει σχήμα πιο πολύπλοκο από αυτό ενός ορθογωνίου.

Στο ακόλουθο θεώρημα έχουμε συγκεντρώσει μερικές από τις πιο χρήσιμες ιδιότητες των διπλών ολοκληρωμάτων. Βεβαιωθείτε ότι τις βρίσκετε όλες διαισθητικά προφανείς.

Θεώρημα 9.1. (Ιδιότητες διπλού ολοκληρώματος) *Έστω $a, b, c \in \mathbb{R}$, συναρτήσεις f, g στο $\mathbb{R}\times\mathbb{R}$, και χωρία $R, S \subseteq \mathbb{R}\times\mathbb{R}$. Το διπλό ολοκλήρωμα ικανοποιεί τις ακόλουθες ιδιότητες:*

1. *(Γραμμικότητα Διπλού Ολοκληρώματος)*

$$\iint_R (af(x,y) + bg(x,y))\, dA = a \iint_R f(x,y)\, dA + b \iint_R g(x,y)\, dA.$$

2. *Αν $f(x,y) \le g(x,y)$ παντού στο R, τότε*

$$\iint_R f(x,y)\, dA \le \iint_R g(x,y)\, dA.$$

3. *Αν* $f(x,y) \geq 0$ *και* $R \subseteq S$, *τότε*

$$\iint_R f(x,y)\,dA \leq \iint_S f(x,y)\,dA.$$

4. *Αν τα χωρία* R *και* S *δεν αλληλοκαλύπτονται, εκτός ίσως στα σύνορά τους, τότε*

$$\iint_{R\cup S} f(x,y)\,dA = \iint_R f(x,y)\,dA + \iint_S f(x,y)\,dA.$$

Εννοείται ότι τα άνω ισχύουν αν υπάρχουν όλα τα ολοκληρώματα που εμφανίζονται.

Θεώρημα 9.2. (Fubini)

1. *Έστω χωρίο* $R \subseteq \mathbb{R}\times\mathbb{R}$ *της μορφής*

$$R = \{(x,y) : a \leq x \leq b,\ \phi_1(x) \leq y \leq \phi_2(x)\}, \tag{9.1}$$

όπου $a, b \in \mathbb{R}$ *με* $a < b$, *και οι συναρτήσεις* $\phi_1(x), \phi_2(x) : [a,b] \to \mathbb{R}$ *είναι συνεχείς με* $\phi_1(x) \leq \phi_2(x)$ *για κάθε* $x \in [a,b]$. *Έστω επίσης* $f(x,y)$ *συνεχής συνάρτηση στο* R. *Το διπλό ολοκλήρωμα της* f *στο* R *υπάρχει και μπορεί να υπολογιστεί μέσω μιας επαλληλίας απλών ολοκληρωμάτων ως εξής:*

$$\iint\limits_R f(x,y)\,dA = \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} f(x,y)\,dy \right) dx. \tag{9.2}$$

2. *Έστω χωρίο* $R \subseteq \mathbb{R}\times\mathbb{R}$ *της μορφής*

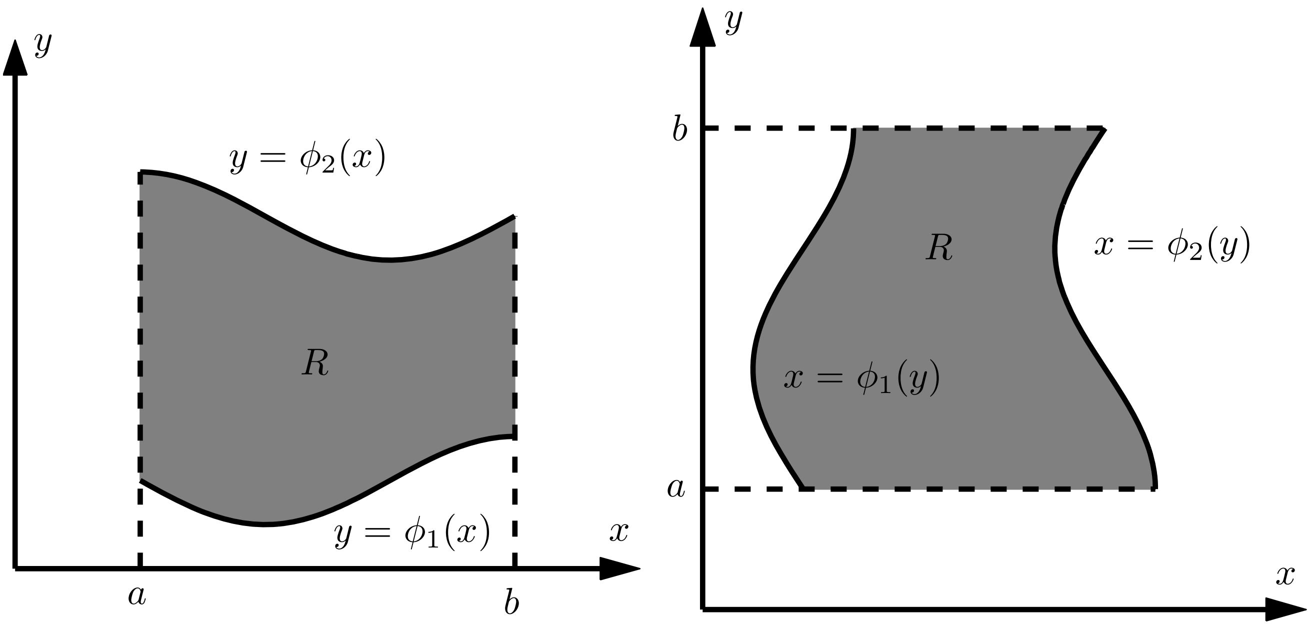
$$R = \{(x,y) : a \leq y \leq b,\ \phi_1(y) \leq x \leq \phi_2(y)\}, \tag{9.3}$$

όπου $a, b \in \mathbb{R}$ *με* $a < b$, *και οι συναρτήσεις* $\phi_1(y), \phi_2(y) : [a,b] \to \mathbb{R}$ *είναι συνεχείς με* $\phi_1(y) \leq \phi_2(y)$ *για κάθε* $y \in [a,b]$. *Έστω επίσης* $f(x,y)$ *συνεχής συνάρτηση στο* R. *Το διπλό ολοκλήρωμα της* f *στο* R *υπάρχει και μπορεί να υπολογιστεί μέσω μιας επαλληλίας απλών ολοκληρωμάτων ως εξής:*

$$\iint\limits_R f(x,y)\,dA = \int_a^b \left(\int_{\phi_1(y)}^{\phi_2(y)} f(x,y)\,dx \right) dy. \tag{9.4}$$

Παρατηρήσεις

1. Το άνω είναι το βασικό εργαλείο για τον υπολογισμό διπλών ολοκληρωμάτων.



Σχήμα 9.2: Τα χωρία των δύο περιπτώσεων του Θεωρήματος Fubini.

2. Οι δύο μορφές του άνω θεωρήματος είναι εντελώς ανάλογες, με τις θέσεις των x, y ανεστραμμένες.

3. Η απόδειξη είναι αρκετά σύνθετη και ξεφεύγει από τους στόχους του μαθήματος.

4. Διαισθητικά, το θεώρημα μας λέει πως για να υπολογίσουμε τον όγκο ενός στερεού με αυθαίρετο σχήμα (το διπλό ολοκλήρωμα) μπορούμε να το κόψουμε σε πολύ λεπτές φέτες, να υπολογίσουμε το εμβαδόν κάθε μιας από αυτές (το εσωτερικό ολοκλήρωμα), και μετά να προσθέσουμε τα εμβαδά αφού τα πολλαπλασιάσουμε με το (πολύ μικρό) πάχος τους (εκτελώντας έτσι την εξωτερική ολοκλήρωση).

5. Συχνά ένα διπλό ολοκλήρωμα μπορεί να υπολογιστεί με χρήση και των δύο μορφών του Θεωρήματος Fubini. Σε πολλές περιπτώσεις, όμως, η μια μορφή οδηγεί σε πολύ απλούστερους υπολογισμούς από την άλλη.

6. Όταν το R είναι ορθογώνιο, δηλαδή $R = [a, b] \times [c, d]$, έχουμε

$$\iint\limits_{[a,b]\times[c,d]} f(x,y)\,dA = \begin{cases} \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx \\ \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy \end{cases}$$

7. Σε εξισώσεις όπου εμφανίζεται επαλληλία απλών ολοκληρωμάτων, όπως στις (9.2), (9.4), μερικές φορές θα παραλείπουμε να γράφουμε τις παρενθέσεις που χωρίζουν την εσωτερική από την εξωτερική ολοκλήρωση. Θα πρέπει πάντα να είμαστε προσεκτικοί στην αντιστοίχιση της κάθε μεταβλητής ολοκλήρωσης με τα σωστά όρια.

8. Συνηθίζεται να γράφουμε και $\iint_R f(x,y)\,dxdy$ αντί για $\iint_R f(x,y)\,dA$. Δεν θα χρησιμοποιήσουμε αυτή τη σύμβαση, για να αποφύγουμε τις παρανοήσεις.

Παράδειγμα 9.1. Σαν ένα απλό παράδειγμα εφαρμογής του Θεωρήματος του Fubini, θα υπολογίσουμε το διπλό ολοκλήρωμα της συνάρτησης $f(x,y) = y^2\sqrt{x}$ σε κάθε ένα από τα ακόλουθα χωρία:

1. Στο ορθογώνιο R_1 με γωνίες τα σημεία $(0,0)$, $(2,0)$, $(0,1)$, και $(2,1)$.

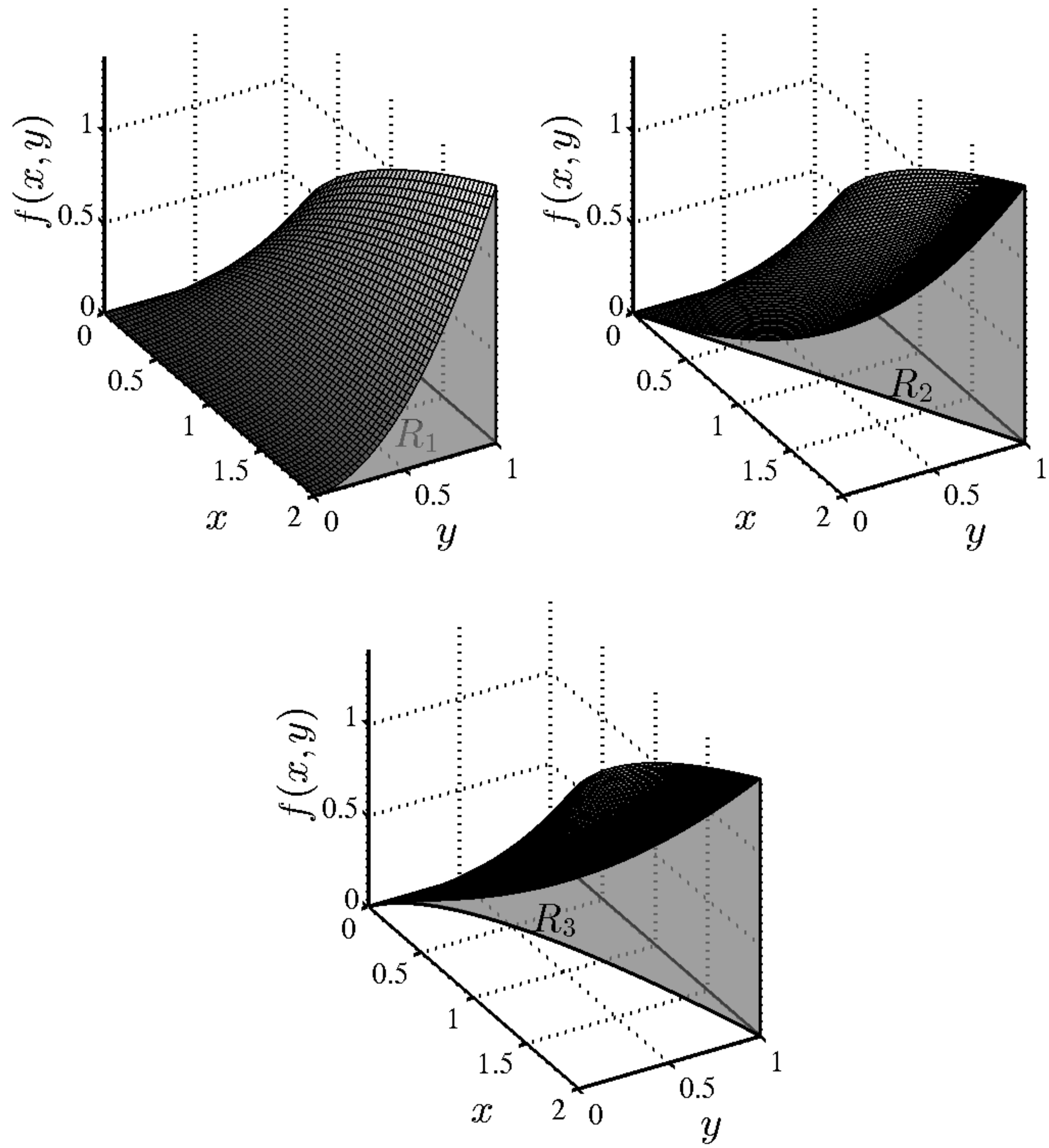
2. Στο τρίγωνο R_2 με γωνίες τα σημεία $(0,0)$, $(0,1)$, και $(2,1)$.

3. Στο χωρίο R_3 που περικλείεται από τις ευθείες $x = 0$, $y = 1$, και την εξίσωση $y = \sqrt{x/2}$.

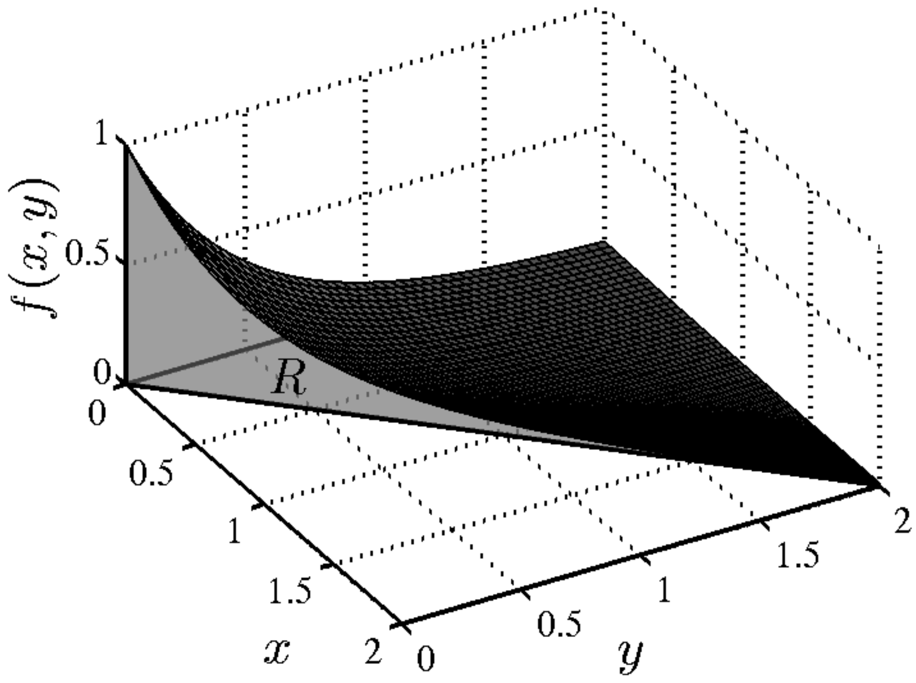
Σε κάθε περίπτωση, το πιο σημαντικό βήμα είναι να περιγράψουμε το χωρίο όπου θα γίνει η ολοκλήρωση σε μια από τις μορφές (9.1), (9.3) που ζητά η εκφώνηση του Θεωρήματος του Fubini. Έχουμε, κατά περίπτωση:

1. Το χωρίο μπορεί να γραφεί ως
$$R_1 = \{(x,y) : 0 \le x \le 2,\ 0 \le y \le 1\},$$
άρα
$$\begin{aligned}\iint_{R_1} f(x,y)\,dA &= \int_0^2 \left(\int_0^1 y^2\sqrt{x}\,dy\right) dx = \int_0^2 \left(\sqrt{x}\int_0^1 \left(\frac{y^3}{3}\right)' dy\right) dx \\ &= \int_0^2 \frac{\sqrt{x}}{3}\,dx = \frac{2}{9}\int_0^2 \left(x^{3/2}\right)' dx = \frac{4\sqrt{2}}{9}.\end{aligned}$$

2. Σε αυτή την περίπτωση,
$$R_2 = \left\{(x,y) : 0 \le x \le 2,\ \frac{x}{2} \le y \le 1\right\},$$
άρα
$$\begin{aligned}\iint_{R_2} f(x,y)\,dA &= \int_0^2 \left(\int_{x/2}^1 y^2\sqrt{x}\,dy\right) dx = \int_0^2 \left(\sqrt{x}\int_{x/2}^1 \left(\frac{y^3}{3}\right)' dy\right) dx \\ &= \int_0^2 \sqrt{x}\left(\frac{1}{3} - \frac{x^3}{24}\right) dx = \int_0^2 \left(\frac{2}{9}x^{3/2} - \frac{1}{108}x^{9/2}\right)' dx \\ &= \frac{4\sqrt{2}}{9} - \frac{16\sqrt{2}}{108} = \frac{8\sqrt{2}}{27}.\end{aligned}$$



Σχήμα 9.3: Τα στερεά του Παραδείγματος 9.1.



Σχήμα 9.4: Το χωρίο και το στερεό του Παραδείγματος 9.2.

3. Σε αυτή την περίπτωση,

$$R_3 = \left\{(x, y) : 0 \leq x \leq 2,\ \sqrt{\frac{x}{2}} \leq y \leq 1\right\},$$

άρα

$$\begin{aligned}
&\iint_{R_3} f(x, y)\, dA \\
&= \int_0^2 \left(\int_{\sqrt{x/2}}^1 y^2\sqrt{x}\, dy\right) dx = \int_0^2 \left(\sqrt{x}\int_{\sqrt{x/2}}^1 \left(\frac{y^3}{3}\right)' dy\right) dx \\
&= \int_0^2 \sqrt{x}\left(\frac{1}{3} - \frac{1}{6\sqrt{2}}x^{3/2}\right) dx = \int_0^2 \left(\frac{2}{9}x^{3/2} - \frac{1}{18\sqrt{2}}x^3\right)' dx \\
&= \frac{4\sqrt{2}}{9} - \frac{8}{18\sqrt{2}} = \frac{2\sqrt{2}}{9}.
\end{aligned}$$

Τα χωρία και τα αντίστοιχα στερεά έχουν σχεδιαστεί στο Σχήμα 9.3. Μπορείτε να υπολογίσετε τα άνω ολοκληρώματα αν χρησιμοποιήσετε την δεύτερη μορφή του Θεωρήματος Fubini;

Παρατήρηση: Το Θεώρημα του Fubini μπορεί να εφαρμοστεί ακόμα και αν το διπλό

ολοκλήρωμα είναι καταχρηστικό, δηλαδή είναι είτε πάνω σε ένα χωρίο που εκτείνεται μέχρι το άπειρο, είτε η ολοκληρωταία συνάρτηση δεν είναι φραγμένη στο χωρίο ή στο όριό του. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 9.2. Θα υπολογίσουμε το ολοκλήρωμα της συνάρτησης $f(x,y) = e^{-x-2y}$ στο τριγωνικό χωρίο που περικλείεται μεταξύ των ευθειών $x = 0$ και $y = x$. Το χωρίο, και το αντίστοιχο στερεό που δημιουργείται, έχουν σχεδιαστεί στο Σχήμα 9.4. Καταρχήν,

$$R = \{(x,y) : 0 \le y \le \infty,\ 0 \le x \le y\}.$$

Παρατηρήστε πως χρησιμοποιούμε την δεύτερη μορφή του Θεωρήματος Fubini, καθώς έτσι καταλήγουμε σε λιγότερες πράξεις. Έχουμε

$$\begin{aligned}
\iint_R f(x,y)\,dA &= \int_0^\infty \left(\int_0^y e^{-x-2y}\,dx\right) dy = \int_0^\infty e^{-2y} \left(\int_0^y e^{-x}\,dx\right) dy \\
&= \int_0^\infty e^{-2y} \left[-e^{-x}\right]_0^y dy = \int_0^\infty e^{-2y}\left(1 - e^{-y}\right) dy \\
&= \int_0^\infty \left(\frac{1}{3}e^{-3y} - \frac{1}{2}e^{-2y}\right)' dy = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.
\end{aligned}$$

9.2 Ζεύγη Συνεχών Τυχαίων Μεταβλητών

Ορισμός 9.1. *Δύο Τ.Μ. καλούνται από κοινού συνεχείς όταν υπάρχει μια συνάρτηση* $f_{XY}(x,y) : \mathbb{R}\times\mathbb{R} \to [0,\infty)$, *η από κοινού συνάρτηση πυκνότητας πιθανότητας, ή από κοινού πυκνότητα, ή απλώς πυκνότητα, τέτοια ώστε για οποιοδήποτε χωρίο* $R \subseteq \mathbb{R}\times\mathbb{R}$, *να ισχύει ότι*

$$P\left[(X,Y)\in R\right] = \iint_R f_{XY}(x,y)\,dA. \tag{9.5}$$

Παρατηρήσεις

1. Στην περίπτωση που έχουμε μια Τ.Μ., οι πιθανότητες που την αφορούν υπολογίζονται μέσω απλών ολοκληρωμάτων της πυκνότητας και ισοδυναμούν με *εμβαδά*. Στην περίπτωση δύο Τ.Μ., οι πιθανότητες υπολογίζονται μέσω διπλών ολοκληρωμάτων και επομένως ισοδυναμούν με *όγκους*! Έστω, για παράδειγμα, πως μας δίνεται η συνάρτηση $f_{XY}(x,y)$ του Σχήματος 9.5(a). Η πιθανότητα το ζεύγος (X,Y) να βρίσκεται σε κάποιο χωρίο R ισούται με τον όγκο του στερεού που οριοθετείται από το χωρίο R και το γράφημα της συνάρτησης. Δείτε για παράδειγμα το Σχήμα 9.5(b).

2. Στην απλούστερη περίπτωση, το χωρίο R είναι κλειστό και φραγμένο, και η συνάρτηση f_{XY} είναι ολοκληρώσιμη σε αυτό με την αυστηρή έννοια του όρου.

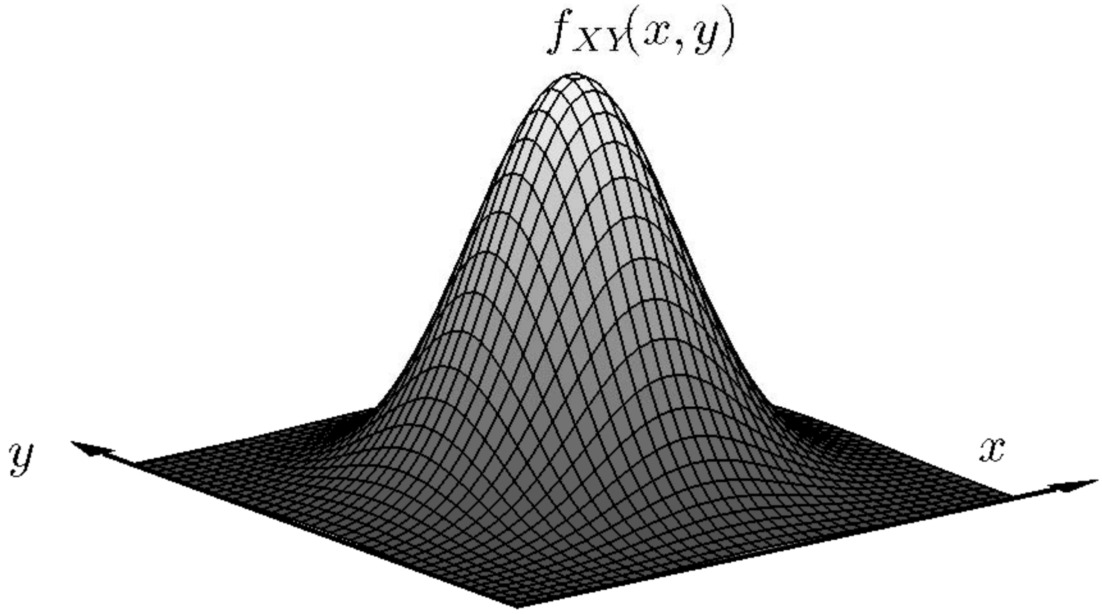
3. Για παράδειγμα, όταν το R είναι καρτεσιανό γινόμενο της μορφής $R=[a,b]\times[c,d]$, η (9.5) γίνεται

$$P\left[(X,Y)\in R\right] = \iint_{[a,b]\times[c,d]} f_{XY}(x,y)\,dA = \begin{cases} \int_a^b \left(\int_c^d f_{XY}(x,y)dy\right)dx, \\ \int_c^d \left(\int_a^b f_{XY}(x,y)dx\right)dy. \end{cases} \tag{9.6}$$

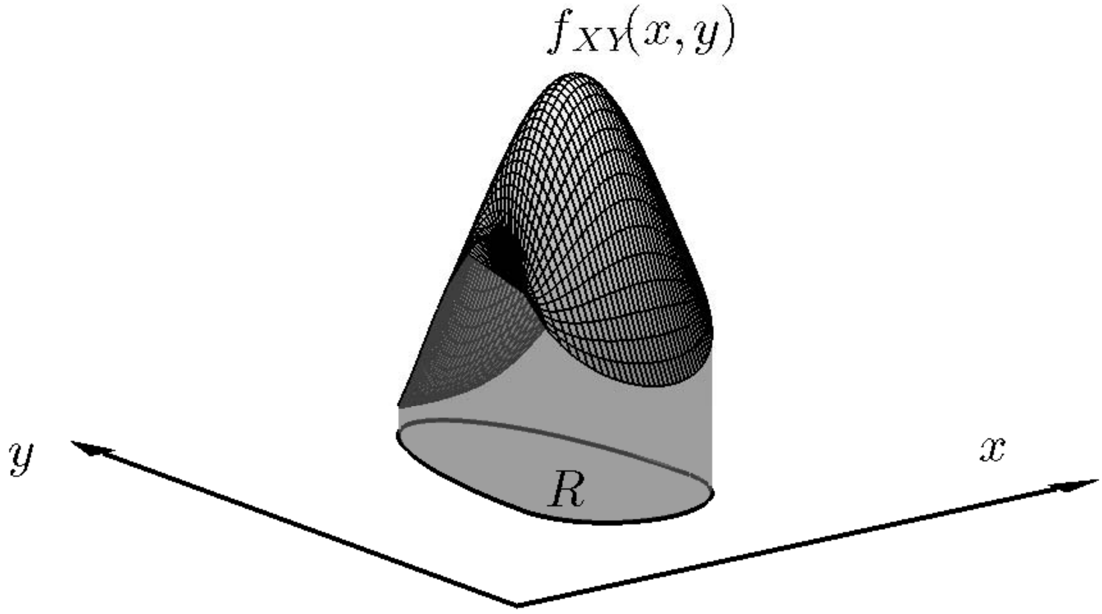
   Παρατηρήστε ότι το διπλό ολοκλήρωμα έχει γραφεί, με δύο διαφορετικούς τρόπους, ως δύο διαδοχικά απλά, με χρήση του Θεωρήματος του Fubini.

4. Στην γενικότερη περίπτωση όμως, το R μπορεί να μην είναι κλειστό (οπότε λαμβάνουμε το ολοκλήρωμα στην κλειστότητά του), ή να είναι ένωση διακριτών χωρίων (οπότε προσθέτουμε τα αντίστοιχα ολοκληρώματα), ή μπορεί το ολοκλήρωμα να υπάρχει μόνο ως καταχρηστικό.

5. Σε κάθε περίπτωση, πρέπει να υπολογίσουμε ένα ή περισσότερα διπλά ολοκληρώματα, πάντα με χρήση του Θεωρήματος του Fubini.

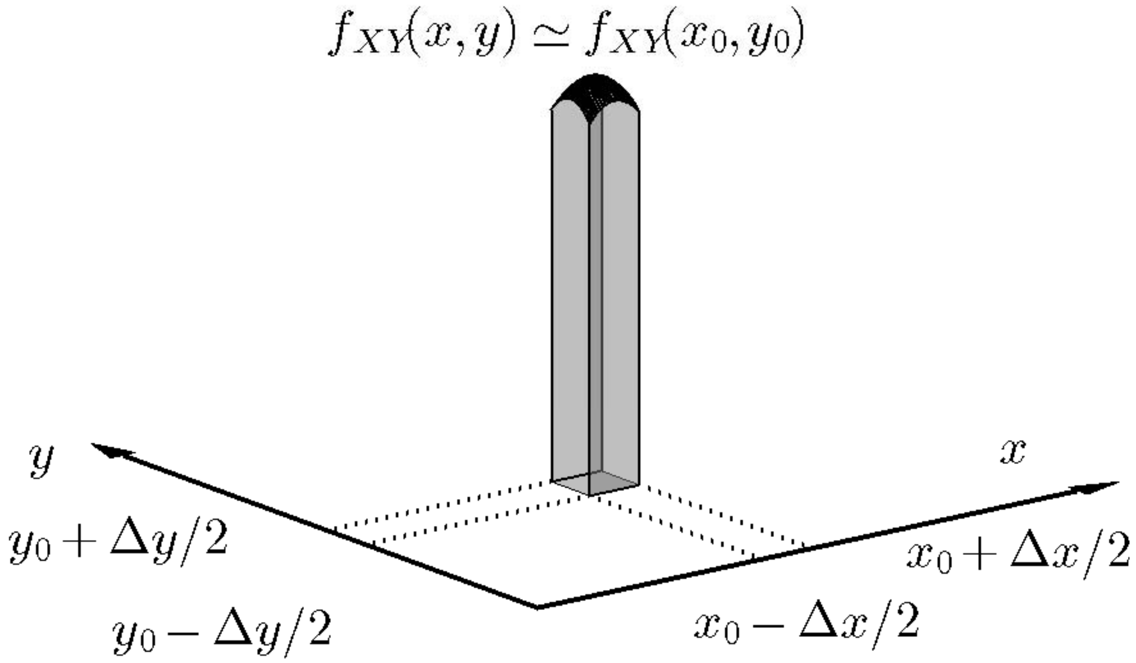


(a)



(b)

Σχήμα 9.5: (a) Μια από κοινού πυκνότητα πιθανότητας $f_{XY}(x, y)$. (b) Η πιθανότητα $P[(X, Y) \in R]$ το ζεύγος (X, Y) να είναι στο χωρίο R δίνεται από τον όγκο του σκιασμένου στερεού που βρίσκεται ανάμεσα στο γράφημα της $f_{XY}(x, y)$ και το R.



Σχήμα 9.6: Καθώς τα $\Delta x, \Delta y \to 0$, το στερεό που δημιουργείται ανάμεσα στο γράφημα της $f_{XY}(x,y)$ και το ορθογώνιο $\left[x_0 - \frac{\Delta x}{2}, x_0 + \frac{\Delta x}{2}\right] \times \left[y_0 - \frac{\Delta y}{2}, y_0 + \frac{\Delta y}{2}\right]$ τείνει σε ένα ορθογώνιο παραλληλεπίπεδο, και ο όγκος του ισούται, προσεγγιστικά, με $f_{XY}(x_0,y_0)\Delta x \Delta y$.

6. Αρκετά συχνά, ο υπολογισμός της πιθανότητας με την μία μέθοδο του Θεωρήματος του Fubini είναι αρκετά απλούστερος από τον υπολογισμό με την άλλη μέθοδο.

7. Διαισθητικά, η από κοινού πυκνότητα στη θέση (x_0, y_0) εκφράζει την πιθανότητα το ζεύγος (X,Y) να έχει τιμές «κοντά» στο (x_0, y_0). Πράγματι, έστω πως η $f(x,y)$ είναι συνεχής στο (x_0,y_0). Τότε

$$\begin{aligned} & P\left(x_0 - \frac{\Delta x}{2} \le X \le x_0 + \frac{\Delta x}{2}, y_0 - \frac{\Delta y}{2} \le Y \le y_0 + \frac{\Delta y}{2}\right) \\ &= P\left((X,Y) \in \left[x_0 - \frac{\Delta x}{2}, x_0 + \frac{\Delta x}{2}\right] \times \left[y_0 - \frac{\Delta y}{2}, y_0 + \frac{\Delta y}{2}\right]\right) \\ &= \iint\limits_{[x_0-\Delta x/2, x_0+\Delta x/2]\times[y_0-\Delta y/2, y_0+\Delta y/2]} f_{XY}(x_0,y_0)\,dy\,dx \simeq f_{XY}(x_0,y_0)\Delta x \Delta y, \end{aligned}$$

όπου το $\simeq$ σημαίνει ότι όταν τα Δx, Δy τείνουν στο 0, τότε το πηλίκο των δύο σκελών αριστερά και δεξιά του $\simeq$ τείνει στη μονάδα. Αυτό προκύπτει από μια γενίκευση του Θεμελιώδους Θεωρήματος του Ολοκληρωτικού Λογισμού στις 2

διαστάσεις, και εξηγείται διαισθητικά στο Σχήμα 9.6. Συνεπώς, όσο πιο μεγάλη είναι η τιμή της πυκνότητας, τόσο πιο πιθανό είναι το ζεύγος (X, Y) να βρεθεί «κοντά» στο (x_0, y_0), αν και βέβαια η πιθανότητα να πάρει ακριβώς την τιμή (x_0, y_0) είναι μηδέν.

8. Όπως και στην περίπτωση της μιας Τ.Μ.:

   (α′) Αν αλλάξουμε την πυκνότητα σε πεπερασμένο πλήθος σημείων, δεν αλλάζει το ολοκλήρωμά της σε οποιοδήποτε χωρίο. Άρα, ένα ζεύγος Τ.Μ. μπορεί να περιγράφεται επακριβώς, όσον αφορά την πιθανότητα να βρίσκεται σε κάποιο χωρίο, από περισσότερες από μια από κοινού πυκνότητες.

   (β′) Για ορισμένα περίπλοκα χωρία R, το ολοκλήρωμα (9.5) δεν ορίζεται, ακόμα και αν η f_{XY} είναι πολύ απλή, π.χ., σταθερή. Αγνοούμε το πρόβλημα.

   (γ′) Οποιαδήποτε συνάρτηση $f_{XY}(x, y) \geq 0$ έχει ολοκλήρωμα

$$\iint\limits_{\mathbb{R}\times\mathbb{R}} f_{XY}(x,y)\,dA = 1$$

   μπορεί να θεωρηθεί πυκνότητα ενός ζεύγους Τ.Μ.

Παράδειγμα 9.3. *(Απλή πυκνότητα πιθανότητας)* Έστω πως το ζεύγος (X, Y) έχει την από κοινού πυκνότητα πιθανότητας

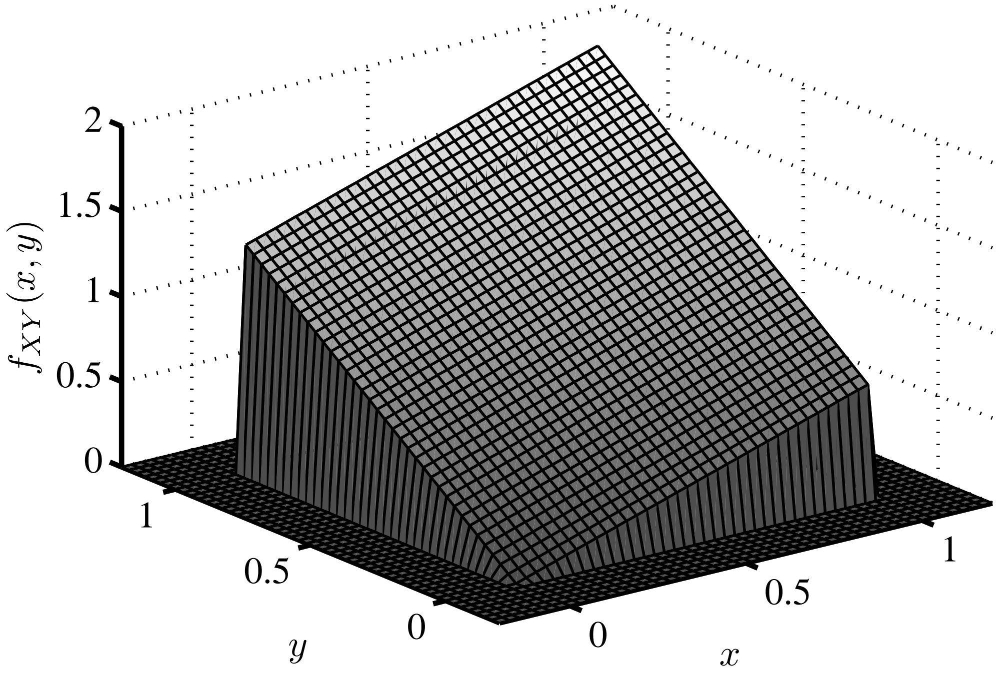
$$f(x,y) = \begin{cases} \frac{1}{3}(2x+4y), & 0 \leq x \leq 1,\ 0 \leq y \leq 1, \\ 0, & \text{αλλού}, \end{cases}$$

η οποία έχει σχεδιαστεί στο Σχήμα 9.7.

1. Θα υπολογίσουμε την πιθανότητα $P\left(0 \leq X \leq \frac{1}{2},\ 0 \leq Y \leq \frac{1}{2}\right)$.
2. Θα υπολογίσουμε την πιθανότητα $P(0 < X < Y)$.

Καταρχήν, παρατηρούμε πως:

$$\begin{aligned}
&\iint\limits_{\mathbb{R}\times\mathbb{R}} f_{XY}(x,y)\,dA \\
&= \iint\limits_{[0,1]\times[0,1]} f_{XY}(x,y)\,dA + \iint\limits_{([0,1]\times[0,1])'} f_{XY}(x,y)\,dA = \iint\limits_{[0,1]\times[0,1]} f_{XY}(x,y)\,dA \\
&= \iint\limits_{[0,1]\times[0,1]} \frac{2x}{3}\,dA + \iint\limits_{[0,1]\times[0,1]} \frac{4y}{3}\,dA = \int_0^1 \frac{2x}{3}\left(\int_0^1 dy\right)dx + \int_0^1 \frac{4y}{3}\left(\int_0^1 dx\right)dy \\
&= \int_0^1 \frac{2x}{3}\,dx + \int_0^1 \frac{4y}{3}\,dy = \int_0^1 \left(\frac{x^2}{3}\right)'dx + \int_0^1 \left(\frac{2y^2}{3}\right)'dy = \frac{1}{3} + \frac{2}{3} = 1.
\end{aligned}$$



Σχήμα 9.7: Η πυκνότητα πιθανότητας του Παραδείγματος 9.3. Εκτός των ορίων $0 \leq x \leq 1$, $0 \leq y \leq 1$, η πυκνότητα είναι 0, ενώ εντός είναι γραμμική ως προς τα x, y.

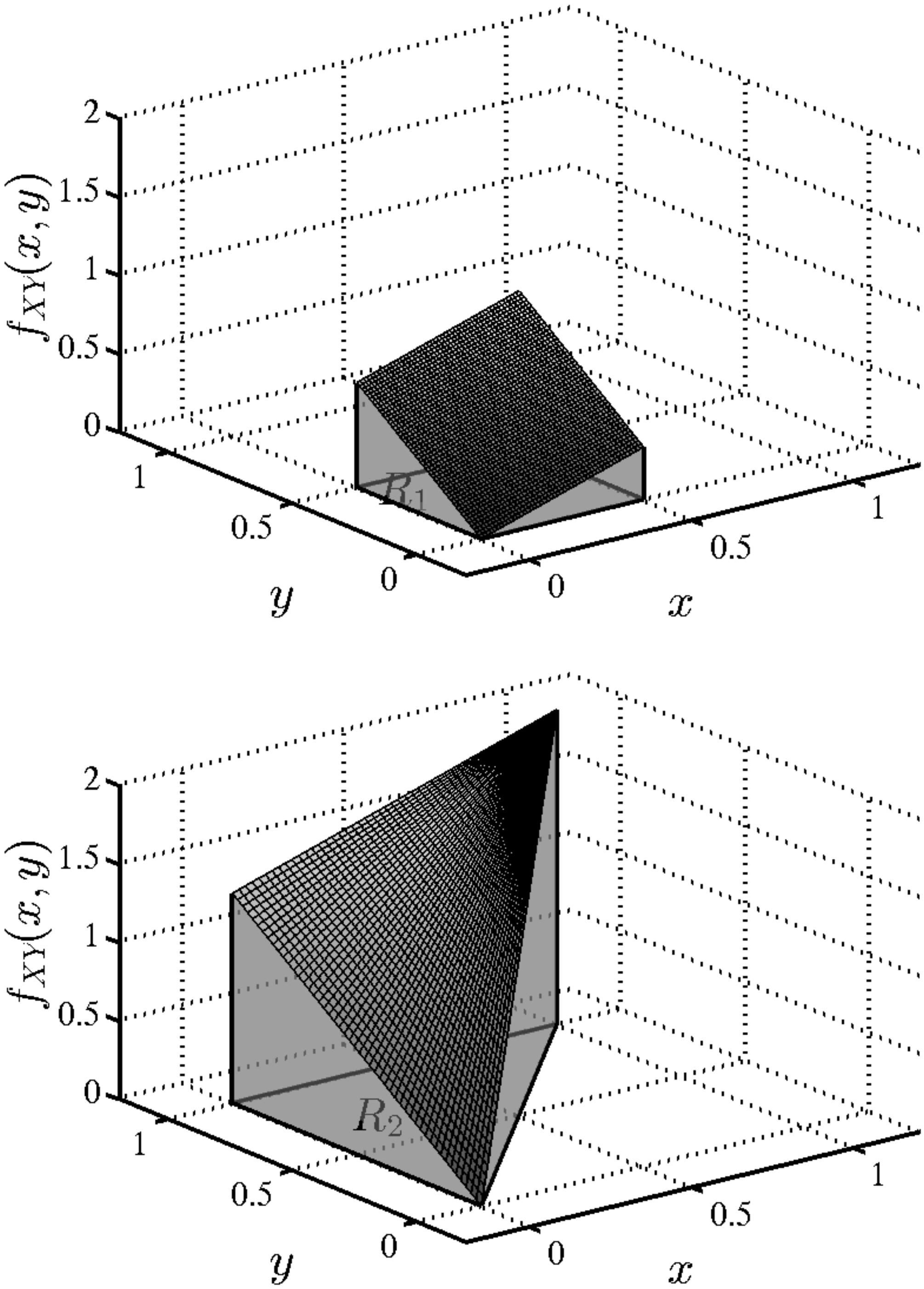
Στην πρώτη ισότητα χρησιμοποιήσαμε το τελευταίο σκέλος του Θεωρήματος 9.1, και στην δεύτερη ισότητα το ότι εκτός του $[0,1] \times [0,1]$ η συνάρτηση είναι 0. Στην τρίτη ισότητα χρησιμοποιήσαμε την γραμμικότητα του διπλού ολοκληρώματος, σύμφωνα με το Θεώρημα 9.1. Στην δεύτερη, εφαρμόσαμε το Θεώρημα του Fubini για κάθε ολοκλήρωμα, βγάζοντας έξω από κάθε εσωτερικό ολοκλήρωμα συντελεστές που δεν εξαρτώνται από την εσωτερική μεταβλητή ολοκλήρωσης. Το τελικό αποτέλεσμα αναμενόταν.

Στην πρώτη περίπτωση, το χωρίο είναι το

$$R_1 = \left\{(x,y) : 0 \leq x \leq \frac{1}{2},\ 0 \leq y \leq \frac{1}{2}\right\},$$

και έχουμε:

$$\begin{aligned}
&P\left(0 \leq X \leq \frac{1}{2},\ 0 \leq Y \leq \frac{1}{2}\right) \\
&= \iint_{[0,\frac{1}{2}]\times[0,\frac{1}{2}]} \left(\frac{2x}{3} + \frac{4y}{3}\right) dA = \int_0^{\frac{1}{2}} \left(\int_0^{\frac{1}{2}} \left(\frac{2x}{3} + \frac{4y}{3}\right) dy \right) dx \\
&= \int_0^{\frac{1}{2}} \left[\frac{2xy}{3} + \frac{2y^2}{3}\right]_0^{\frac{1}{2}} dx = \int_0^{\frac{1}{2}} \left(\frac{x}{3} + \frac{1}{6}\right) dx = \int_0^{\frac{1}{2}} \left(\frac{x^2}{6} + \frac{x}{6}\right)' dx = \frac{1}{8}.
\end{aligned}$$



Σχήμα 9.8: Οι όγκοι των στερεών ισούνται με τις πιθανότητες $P\left(0 \leq X \leq \frac{1}{2},\ 0 \leq Y \leq \frac{1}{2}\right)$ και $P[(X < Y)$ του Παραδείγματος 9.3.

Παρατηρήστε ότι εδώ επιλέξαμε να μην χρησιμοποιήσουμε την γραμμικότητα του ολοκληρώματος, και έτσι εφαρμόσαμε το Θεώρημα του Fubini μια φορά. Το στερεό του οποίου τον όγκο υπολογίσαμε έχει σχεδιαστεί στο Σχήμα 9.8.

Στην δεύτερη περίπτωση, το χωρίο έχει τριγωνικό σχήμα, εκτείνεται στο άπειρο, και μπορεί να περιγραφεί ως

$$R = \{(x,y) : 0 \le y < \infty,\ 0 \le x \le y\}.$$

Όμως, όπως και στην προηγούμενη περίπτωση, η συνάρτηση είναι μηδενική εκτός του τετραγώνου $[0,1] \times [0,1]$, άρα τελικά αρκεί να ολοκληρώσουμε στην τομή των δύο χωριών, που περιγράφεται ως

$$R_2 = \{(x,y) : 0 \le y \le 1,\ 0 \le x \le y\}.$$

Κατά τα γνωστά από το Θεώρημα Fubini, έχουμε:

$$\begin{aligned} P(X<Y) &= \int_0^1 \left(\int_0^y \left(\frac{2x}{3} + \frac{4y}{3} \right) dx \right) dy = \int_0^1 \left(\int_0^y \left(\frac{x^2}{3} + \frac{4yx}{3} \right)' dx \right) dy \\ &= \int_0^1 \left(\frac{y^2}{3} + \frac{4y^2}{3} \right) dy = \frac{5}{3} \int_0^1 y^2\, dy = \frac{5}{9} \int_0^1 (y^3)'\, dy = \frac{5}{9}. \end{aligned}$$

Το στερεό του οποίου τον όγκο υπολογίσαμε φαίνεται σκιασμένο στο Σχήμα 9.8.

Παράδειγμα 9.4. *(Πυκνότητα πιθανότητας που εκτείνεται στο άπειρο)* Έστω πως το ζεύγος (X,Y) έχει την από κοινού πυκνότητα πιθανότητας

$$f(x,y) = \frac{1}{2} e^{-2|x|-|y|}, \quad x, y \in \mathbb{R},$$

η οποία έχει σχεδιαστεί στο Σχήμα 9.9. Όπως και στο προηγούμενο παράδειγμα,

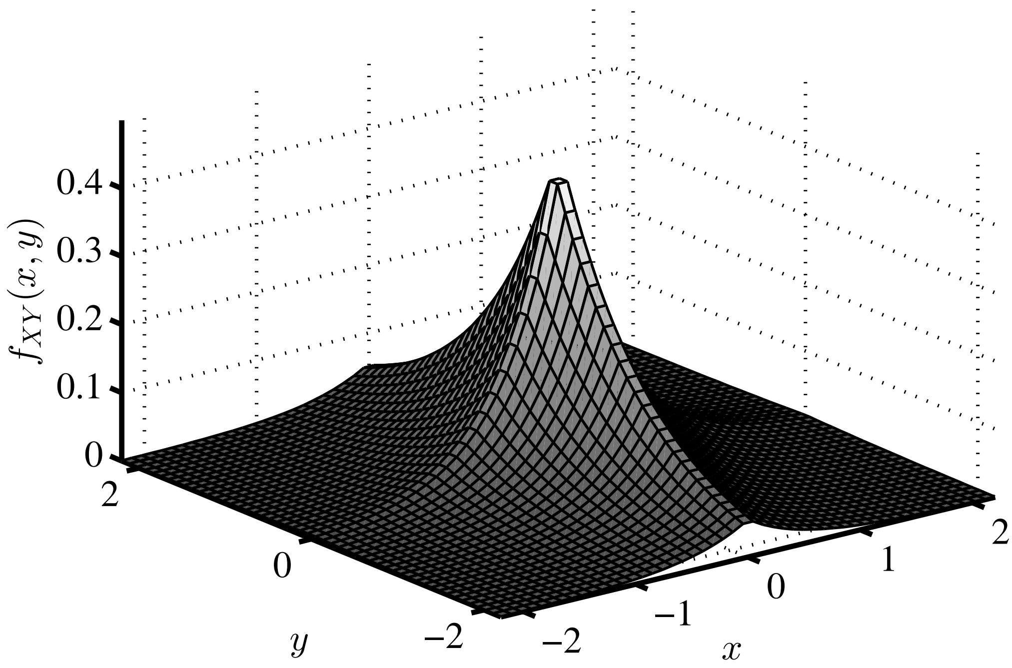
1. Θα υπολογίσουμε την πιθανότητα $P\left(0 \le X \le \frac{1}{2},\ 0 \le Y \le \frac{1}{2}\right)$.
2. Θα υπολογίσουμε την πιθανότητα $P(0 < X < Y)$.

Καταρχήν, παρατηρούμε πως:

$$\begin{aligned} \iint\limits_{\mathbb{R}\times\mathbb{R}} f_{XY}(x,y)\, dA &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \frac{1}{2} e^{-2|x|-|y|} \right) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{2} e^{-2|x|} \left(\int_{-\infty}^{\infty} e^{-|y|}\, dy \right) dx = \int_{-\infty}^{\infty} e^{-2|x|}\, dx = 1. \end{aligned}$$

Στην τρίτη ισότητα χρησιμοποιήσαμε το ότι

$$\int_{-\infty}^{\infty} e^{-|y|}\, dy = 2 \int_0^{\infty} e^{-y}\, dy = 2\left[-e^{-y}\right]_0^{\infty} = 2, \tag{9.7}$$



Σχήμα 9.9: Η πυκνότητα πιθανότητας του Παραδείγματος 9.4. Η πυκνότητα είναι θετική παντού στο $\mathbb{R} \times \mathbb{R}$.

ενώ στην τέταρτη ισότητα το ότι

$$\int_{-\infty}^{\infty} e^{-2|x|}\,dx = 2\int_0^{\infty} e^{-2x}\,dx = \left[-e^{-2x}\right]_0^{\infty} = 1. \tag{9.8}$$

Το τελικό αποτέλεσμα αναμενόταν, καθώς η πιθανότητα να πάρει το ζεύγος (X, Y) οποιαδήποτε τιμή πρέπει να είναι μονάδα. Στην πρώτη περίπτωση, το χωρίο είναι το

$$R_1 = \left\{(x,y) : 0 \le x \le \frac{1}{2},\ 0 \le y \le \frac{1}{2}\right\},$$

και έχουμε:

$$\begin{aligned}
&P\left(0 \le X \le \frac{1}{2},\ 0 \le Y \le \frac{1}{2}\right)\\
&= \iint_{[0,\frac{1}{2}]\times[0,\frac{1}{2}]} \left(\frac{e^{-2x-y}}{2}\right) dA = \int_0^{\frac{1}{2}} \left(\int_0^{\frac{1}{2}} \frac{e^{-2x-y}}{2}\,dy\right) dx = \int_0^{\frac{1}{2}} \frac{e^{-2x}}{2} \left(\int_0^{\frac{1}{2}} e^{-y}\,dy\right) dx\\
&= \int_0^{\frac{1}{2}} \frac{1}{2}e^{-2x}\left[-e^{-y}\right]_0^{\frac{1}{2}} dx = \int_0^{\frac{1}{2}} \frac{1}{2}e^{-2x}\left(1-e^{-1/2}\right) dx = \frac{1-e^{-1/2}}{2}\int_0^{\frac{1}{2}} e^{-2x}\,dx\\
&= \frac{1-e^{-1}}{2}\left[-\frac{1}{2}e^{-2x}\right]_0^{\frac{1}{2}} = \frac{1}{4}(1-e^{-1})(1-e^{-1/2}).
\end{aligned}$$

Το στερεό του οποίου τον όγκο υπολογίσαμε έχει σχεδιαστεί στο Σχήμα 9.10.

Στην δεύτερη περίπτωση, το χωρίο μπορεί να περιγραφεί ως εξής:

$$R_2 = \{(x, y) : 0 \le y \le \infty,\ 0 \le x \le y\},$$

Κατά τα γνωστά από το Θεώρημα Fubini, έχουμε:

$$\begin{aligned} P(0 < X < Y) &= \int_0^\infty \left(\int_0^y \frac{1}{2} e^{-2x-y}\, dx \right) dy = \int_0^\infty \frac{1}{2} e^{-y} \left(\int_0^y e^{-2x}\, dx \right) dy \\ &= \int_0^\infty \frac{1}{2} e^{-y} \left[-\frac{e^{-2x}}{2} \right]_0^y dy = \int_0^\infty \frac{1}{2} e^{-y} \left(\frac{1}{2} - e^{-2y} \right) dy \\ &= \int_0^\infty \left(\frac{1}{4} e^{-y} - \frac{1}{2} e^{-3y} \right) dy = \left[-\frac{1}{4} e^{-y} + \frac{1}{6} e^{-3y} \right]_0^\infty = \frac{1}{4} - \frac{1}{6} = \frac{1}{12}. \end{aligned}$$

Το στερεό του οποίου τον όγκο υπολογίσαμε φαίνεται σκιασμένο στο Σχήμα 9.10.

Λήμμα 9.1. (Βασικές ιδιότητες της από κοινού πυκνότητας) *Έστω από κοινού συνεχείς Τ.Μ. με από κοινού πυκνότητα $f_{XY}(x, y)$. Ισχύουν τα ακόλουθα:*

1. $\displaystyle\iint_{\mathbb{R}\times\mathbb{R}} f(x)\, dA = 1.$

2. *Το ορθογώνιο χωρίο R με πλευρές $x = a$, $x = b$, $y = c$, $y = d$, όπου $a \le b$, $c \le d$, έχει πιθανότητα*

$$P((X, Y) \in R) = \iint_{[a,b]\times[c,d]} f_{XY}(x, y)\, dA = \begin{cases} \int_a^b \left(\int_c^d f_{XY}(x, y) dy \right) dx, \\ \int_c^d \left(\int_a^b f_{XY}(x, y) dx \right) dy, \end{cases}$$

   ανεξάρτητα από το αν και πόσα από τα οριακά του σημεία ανήκουν σε αυτό ή όχι.

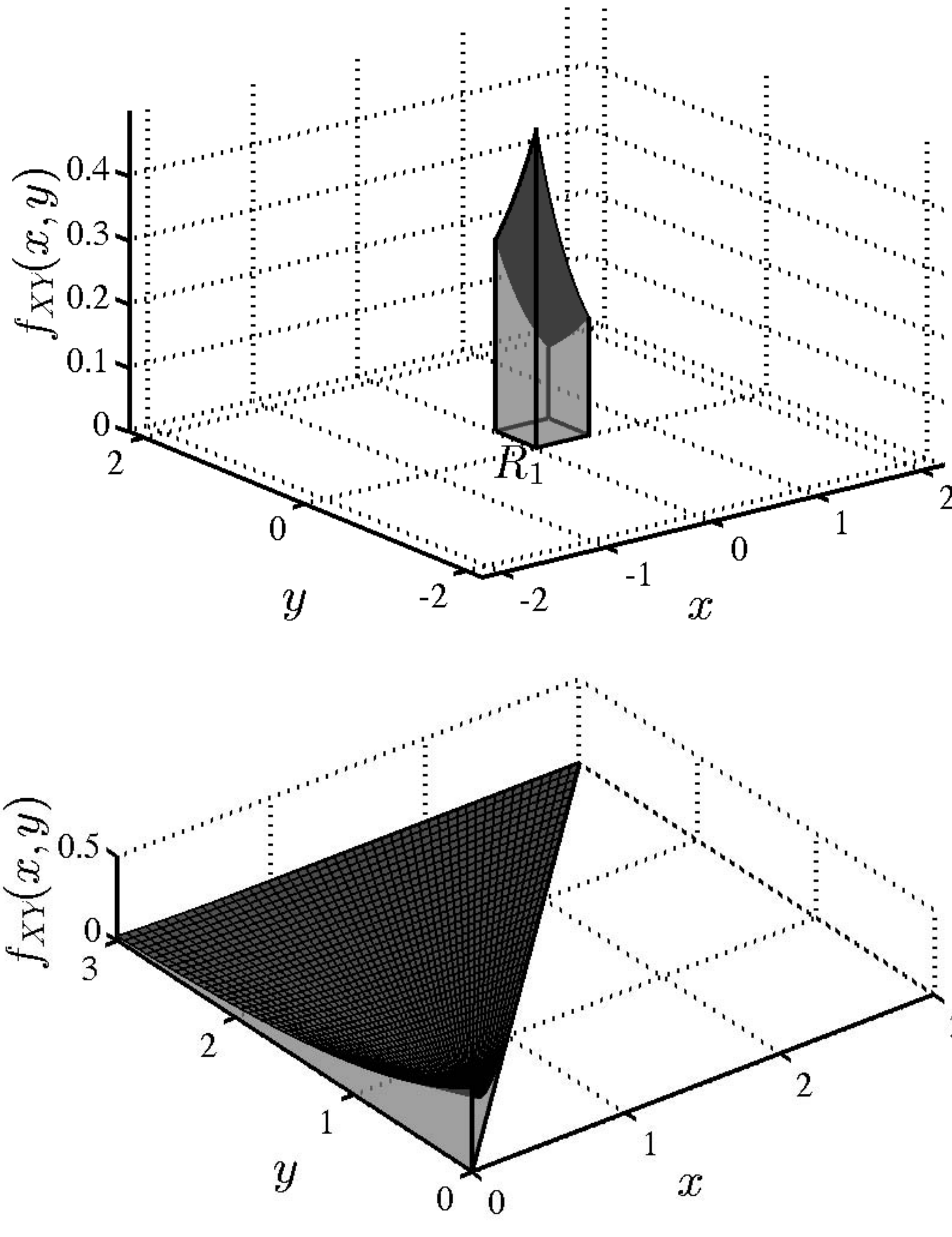
3. $P(X = a, Y = b) = 0$ *για οποιοδήποτε ζεύγος $a, b \in \mathbb{R}$.*

4. *Η X είναι συνεχής με πυκνότητα*

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy. \tag{9.9}$$

5. *Η Y είναι συνεχής με πυκνότητα*

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx. \tag{9.10}$$



Σχήμα 9.10: Οι όγκοι των στερεών ισούνται με τις πιθανότητες $P\left(0 \le X \le \frac{1}{2},\ 0 \le Y \le \frac{1}{2}\right)$ και $P(X < Y)$ του Παραδείγματος 9.3.

Απόδειξη. 1. Προκύπτει από το γεγονός ότι η πιθανότητα να πάρει το ζεύγος (X, Y) οποιαδήποτε τιμή, που ισούται με το δοσμένο ολοκλήρωμα, πρέπει να είναι μονάδα.

2. Προκύπτει από τον ορισμό (9.5).

3. Προκύπτει από τον ορισμό (9.5).

4. Παρατηρήστε πως, για οποιοδήποτε $A \subseteq \mathbb{R}$,

$$P(X \in A) = P(X \in A, Y \in \mathbb{R}) = \iint_{A\times\mathbb{R}} f_{XY}(x, y)\, dA$$
$$= \int_A \left(\int_{\mathbb{R}} f_{XY}(x, y)\, dy \right) dx \Rightarrow P(X \in A) = \int_A \left(\int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \right) dx.$$

Άρα, βρήκαμε μια συνάρτηση $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy$ η οποία λαμβάνει μη αρνητικές τιμές και η οποία επιπλέον έχει την ιδιότητα, για οποιοδήποτε $A \subseteq \mathbb{R}$,

$$P(X \in A) = \int_A f_X(x)\, dx.$$

Άρα εξ ορισμού των συνεχών Τ.Μ., η X είναι συνεχής, με πυκνότητα που δίνεται από την (9.9).

5. Προκύπτει εντελώς ανάλογα με την προηγούμενη περίπτωση.

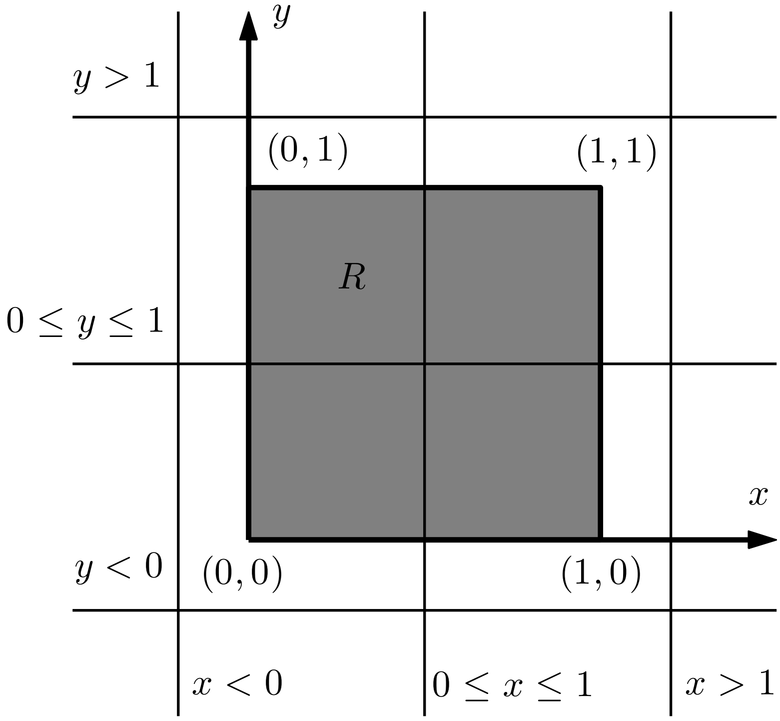
□

Ορισμός 9.2. *Αν οι X, Y είναι από κοινού συνεχείς, οι πυκνότητές τους $f_X(x)$, $f_Y(y)$ καλούνται περιθώριες πυκνότητες πιθανότητας.*

Παράδειγμα 9.5. Θα υπολογίσουμε τις περιθώριες πυκνότητες του Παραδείγματος 9.4. Με εφαρμογή των (9.9), (9.10), έχουμε:

$$f_X(x) = \int_{-\infty}^{\infty} \frac{e^{-2|x|-|y|}}{2}\, dy = \frac{e^{-2|x|}}{2} \int_{-\infty}^{\infty} e^{-|y|}\, dy = e^{-2|x|},\ x \in \mathbb{R},$$
$$f_Y(y) = \int_{-\infty}^{\infty} \frac{e^{-2|x|-|y|}}{2}\, dx = \frac{e^{-|y|}}{2} \int_{-\infty}^{\infty} e^{-2|x|}\, dx = \frac{e^{-|y|}}{2},\ y \in \mathbb{R}.$$

Στις άνω, χρησιμοποιήσαμε τις (9.7), (9.8). Η περίπτωση αυτή είναι αρκετά απλή, γιατί η από κοινού πυκνότητα δεν είναι κλαδική, κάτι που δεν συμβαίνει συχνά. Δείτε το ακόλουθο παράδειγμα για μια κάπως πιο σύνθετη περίπτωση.



Σχήμα 9.11: Παράδειγμα 9.6: Η από κοινού πυκνότητα είναι μη μηδενική στο χωρίο $R = [0,1] \times [0,1]$. Για να υπολογίσουμε τις περιθώριες πυκνότητες $f_X(x)$, $f_Y(y)$, πρέπει να ολοκληρώσουμε την από κοινού κατά μήκος ευθειών καθέτων ή οριζοντίων (αντίστοιχα), έχοντας την μεταβλητή που δεν ολοκληρώνεται, x ή y αντίστοιχα, ως μια παράμετρο που θα επηρεάσει την τιμή του ολοκληρώματος μέσω της θέσης της ευθείας.

Παράδειγμα 9.6. Θα υπολογίσουμε τις περιθώριες πυκνότητες του Παραδείγματος 9.3. Για να υπολογίσουμε την $f_X(x)$, παρατηρούμε πως αν $x < 0$ ή $x > 1$, τότε προκύπτει ότι $f_X(x) = 0$. Πράγματι, σε αυτή την περίπτωση η $f_X(x)$ ισούται με το ολοκλήρωμα μιας συνάρτησης που είναι παντού 0. Αν $0 \leq x \leq 1$, τότε

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_0^1 \left(\frac{2x}{3} + \frac{4y}{3}\right) dy \\ &= \frac{2x}{3}\int_0^1 dy + \frac{4}{6}\int_0^1 \left(y^2\right)'\,dy = \frac{2}{3}(x+1). \end{aligned}$$

Δείτε το Σχήμα 9.11 για τις τρεις περιπτώσεις. Συγκεντρωτικά:

$$f_X(x) = \begin{cases} \frac{2}{3}(x+1), & x \in [0,1], \\ 0, & x \notin [0,1]. \end{cases}$$

Παρομοίως,

$$f_Y(y) = \begin{cases} \frac{1}{3}(4y+1), & y \in [0,1], \\ 0, & y \notin [0,1]. \end{cases}$$

9.3 Μέση Τιμή και Συνδιακύμανση

Λήμμα 9.2. (Μέση τιμή συνάρτησης δύο Τ.Μ.) *Έστω δύο συνεχείς Τ.Μ. X, Y, με από κοινού πυκνότητα πιθανότητας $f_{XY}(x,y)$.*

1. *Έστω Τ.Μ. $Z = g(X,Y)$. Η μέση τιμή της ισούται με*

$$E(Z) = \iint_{\mathbb{R}\times\mathbb{R}} g(x,y) f_{XY}(x,y)\, dA = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)\, dxdy.$$

2. *$E(aX + bY + c) = aE(X) + bE(Y) + c$.*

3. *Πιο γενικά, αν έχουμε K Τ.Μ. $Z_k = g_k(X,Y)$, $k = 1,\ldots,K$, τότε*

$$E\left(\sum_{k=1}^{K} g_k(X,Y)\right) = \sum_{k=1}^{K} E(g_k(X,Y)).$$

Απόδειξη. 1. Η απόδειξη είναι πολύπλοκη και παραλείπεται.

2. Παρατηρούμε πως:

$$\begin{aligned}
&E(aX + bY + c)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax + by + c) f_{XY}(x,y)\, dxdy\\
&= a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{XY}(x,y)\, dxdy + b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x,y)\, dxdy\\
&\quad + c\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\, dxdy\\
&= a\int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f_{XY}(x,y)\, dy\right) dx + b\int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} f_{XY}(x,y)\, dx\right) dy + c\\
&= a\int_{-\infty}^{\infty} x f_X(x)\, dx + b\int_{-\infty}^{\infty} y f_Y(y)\, dy + c\\
&= aE(X) + bE(Y) + c.
\end{aligned}$$

Στην πρώτη ισότητα χρησιμοποιήσαμε το πρώτο σκέλος. Στην δεύτερη, τη γραμμικότητα του διπλού ολοκληρώματος. Στην τρίτη, το ότι το ολοκλήρωμα της από κοινού πυκνότητας σε όλο το επίπεδο πρέπει να ισούται με μονάδα. Στην τέταρτη, τις (9.9) και (9.10). Στην πέμπτη, τον ορισμό της μέσης τιμής για τις συναρτήσεις μίας μεταβλητής. Παρατηρήστε ότι η ιδιότητα θα μπορούσε εναλλακτικά να είχε δειχθεί αν επικαλούμασταν το πρώτο σκέλος στο τέλος της δεύτερης ισότητας.

3.
$$\begin{aligned} E\left(\sum_{k=1}^{K} g_k(X,Y)\right) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\sum_{k=1}^{K} g_k(X,Y)\right) f_{XY}(x,y)\,dxdy \\ &= \sum_{k=1}^{K}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_k(X,Y) f_{XY}(x,y)\,dxdy \\ &= \sum_{k=1}^{K} E(g_k(X,Y)). \end{aligned}$$

Στην δεύτερη ισότητα, χρησιμοποιήσαμε την γραμμικότητα του διπλού ολοκληρώματος. Στην τρίτη, το πρώτο σκέλος.

□

Παρατήρηση: Το λήμμα μας επιτρέπει να υπολογίσουμε μέσες τιμές τυχαίων μεταβλητών με πολύ λιγότερο κόπο σε σχέση με την περίπτωση που χρησιμοποιούσαμε τον ορισμό της μέσης τιμής. Δείτε το ακόλουθο παράδειγμα.

Παράδειγμα 9.7. Για τις Τ.Μ. X, Y του Παραδείγματος 9.3 έχουμε

$$\begin{aligned} E(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{XY}(x,y)\,dxdy = \frac{1}{3}\int_0^1\left(\int_0^1 xy(2x+4y)\,dy\right)dx \\ &= \frac{1}{3}\int_0^1\left(\int_0^1\left(x^2y^2+\frac{4}{3}xy^3\right)' dy\right)dx = \frac{1}{3}\int_0^1\left(x^2+\frac{4x}{3}\right)dx \\ &= \frac{1}{3}\int_0^1\left(\frac{x^3}{3}+\frac{2x^2}{3}\right)' dx = \frac{1}{3}\left(\frac{1}{3}+\frac{2}{3}\right) = \frac{1}{3}. \end{aligned}$$

Χωρίς χρήση του άνω λήμματος, για να υπολογίσουμε την μέση τιμή $E(Z)$ της Τ.Μ. $Z \triangleq XY$, θα έπρεπε

1. να υπολογίσουμε την κατανομή $F_Z(z) = P(Z \le z) = P(XY \le z)$, μέσω του διπλού ολοκληρώματος
$$P(XY \le z) = \iint_{xy\le z} f_{XY}(x,y)\,dA = \iint_{xy\le z,\ 0\le x,y\le 1}\left(\frac{2x}{3}+\frac{4y}{3}\right)dA.$$
2. να υπολογίσουμε την πυκνότητα της Z, με χρήση της σχέσης $f_Z(z) = F_Z'(z)$, και
3. κατόπιν να εφαρμόσουμε τον ορισμό της μέσης τιμής $E(Z) = \int_{-\infty}^{\infty} z f_Z(z)\,dz$.

Η διαδικασία θα ήταν πολύ πιο μακροσκελής (ιδιαίτερα το πρώτο βήμα).

Ορισμός 9.3. *Η <u>συνδιακύμανση</u>* $\mathrm{COV}(X,Y)$ *μεταξύ δύο συνεχών Τ.Μ.* X,Y *ορίζεται ως:*

$$\mathrm{COV}(X,Y) \triangleq E\big[\big(X - E(X)\big)\big(Y - E(Y)\big)\big].$$

Παρατήρηση: Μια πρώτη διαισθητική ερμηνεία της συνδιακύμανσης είναι πως, όταν $\mathrm{COV}(X,Y) > 0$, τότε οι δύο Τ.Μ. X,Y τείνουν να παίρνουν και οι δύο ταυτόχρονα «μεγάλες» τιμές, ή και οι δύο ταυτόχρονα «μικρές» τιμές. Αντίστοιχα, όταν $\mathrm{COV}(X,Y) < 0$, τότε όταν η μία Τ.Μ. παίρνει μεγάλες τιμές η άλλη τείνει να παίρνει μικρές τιμές. Άρα η συνδιακύμανση $\mathrm{COV}(X,Y)$ παρέχει μια πρώτη ένδειξη για τη σχέση ανάμεσα στις X,Y.

Λήμμα 9.3. (Ιδιότητες συνδιακύμανσης) *Έστω* X,Y *συνεχείς Τ.Μ.*

1. $\mathrm{COV}(X,X) = \mathrm{VAR}(X)$.
2. $\mathrm{COV}(X,-X) = -\mathrm{VAR}(X)$.
3. $$\mathrm{COV}(X,Y) = E(XY) - E(X)E(Y). \tag{9.11}$$
4. $$\mathrm{VAR}(X+Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\mathrm{COV}(X,Y). \tag{9.12}$$

Απόδειξη. Η απόδειξη είναι ίδια με την απόδειξη για την διακριτή περίπτωση, και συνεπώς παραλείπεται. □

Παράδειγμα 9.8. Για τις Τ.Μ. X, Y του Παραδείγματος 9.3 έχουμε, χρησιμοποιώντας τις περιθώριες πυκνότητες που βρήκαμε στο Παράδειγμα 9.6:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^1 x\frac{2}{3}(x+1)\,dx = \int_0^1 \left(\frac{2x^3}{9} + \frac{x^2}{3}\right)' dx = \frac{2}{9} + \frac{1}{3} = \frac{5}{9}.$$

Παρομοίως,

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y f_Y(y)\,dy = \int_0^1 y\frac{1}{3}(4y+1)\,dy \\ &= \int_0^1 \frac{1}{3}\left(\frac{4y^3}{3} + \frac{y^2}{2}\right)' dy = \frac{1}{3}\left(\frac{4}{3} + \frac{1}{2}\right) = \frac{11}{18}. \end{aligned}$$

Τέλος, στο Παράδειγμα 9.7 βρήκαμε πως $E(XY) = \frac{1}{3}$, άρα τελικά

$$\mathrm{COV}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{5}{9} \times \frac{11}{18} = -\frac{1}{162}.$$

9.4 Ανεξάρτητες Συνεχείς Τυχαίες Μεταβλητές

Ορισμός 9.4. (Ζεύγη ανεξάρτητων συνεχών Τ.Μ.) *Δύο συνεχείς Τ.Μ. X, Y καλούνται ανεξάρτητες αν για οποιαδήποτε υποσύνολα $A, B \subseteq \mathbb{R}$ ισχύει*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B). \tag{9.13}$$

Παρατήρηση: Όπως και στην περίπτωση των διακριτών Τ.Μ., δύο Τ.Μ. είναι ανεξάρτητες αν δύο οποιαδήποτε ενδεχόμενα που αφορούν το καθένα αποκλειστικά τη μία από τις δύο Τ.Μ. είναι ανεξάρτητα. Η ανεξαρτησία λοιπόν και των συνεχών Τ.Μ. είναι άμεσα σχετισμένη με την ανεξαρτησία των ενδεχόμενων.

Λήμμα 9.4. (Κριτήριο ανεξαρτησίας Τ.Μ.)

1. *Δύο συνεχείς Τ. Μ. X, Y, με από κοινού πυκνότητα $f_{XY}(x, y)$ και περιθώριες πυκνότητες $f_X(x)$, $f_Y(y)$ είναι ανεξάρτητες αν*

$$f_{XY}(x, y) = f_X(x)f_Y(y), \quad \forall x, y \in \mathbb{R}. \tag{9.14}$$

2. *Αν δύο Τ.Μ. X, Y, με πυκνότητες, αντιστοίχως, $f_X(x)$, $f_Y(y)$, είναι ανεξάρτητες, τότε μπορούμε να θέσουμε ως από κοινού πυκνότητά τους την $f_{XY}(x, y) = f_X(x)f_Y(y)$.*

Απόδειξη. Θα αποδείξουμε μόνο το πρώτο σκέλος. Έστω πως ισχύει η (9.14) και έστω δύο ενδεχόμενα $A, B \subseteq \mathbb{R}$. Παρατηρούμε πως

$$\begin{aligned} P(X \in A, Y \in B) &= \iint_{A \times B} f_{XY}(x, y)\, dA = \int_A \left(\int_B f_X(x) f_Y(y)\, dy \right) dx \\ &= \int_A f_X(x) \left(\int_B f_Y(y)\, dy \right) dx \\ &= \left(\int_B f_Y(y)\, dy \right) \left(\int_A f_X(x)\, dx \right) = P(X \in A)P(Y \in B). \end{aligned}$$

Η δεύτερη ισότητα προέκυψε από εφαρμογή του Θεωρήματος του Fubini (σε μια πιο γενική μορφή από αυτή που έχουμε δει) και την υπόθεση, ενώ η τρίτη και η τέταρτη βγάζοντας σταθερές εκτός ολοκληρωμάτων. □

Παρατήρηση: Μπορεί δύο Τ.Μ. να είναι ανεξάρτητες χωρίς η από κοινού τους πυκνότητα να ικανοποιεί την (9.14). Για να καταλάβετε γιατί, σκεφτείτε ως εξής: Έστω

δύο Τ.Μ. X, Y για τις οποίες η (9.14) ικανοποιείται, οπότε οι Τ.Μ. είναι ανεξάρτητες. Αν αλλάξουμε την από κοινού πυκνότητα σε ένα μόνο σημείο, τότε το κριτήριο δεν ικανοποιείται πλέον, όμως οι X, Y συνεχίζουν να είναι ανεξάρτητες, αφού η αλλαγή της από κοινού πυκνότητας σε ένα σημείο δεν μπορεί να επηρεάσει τον υπολογισμό του ολοκληρώματος της σε κανένα σύνολο, άρα και τις τιμές των πιθανοτήτων που εμφανίζονται στα δύο σκέλη της (9.13).

Παράδειγμα 9.9. Θα εξετάσουμε κατά πόσον οι X, Y του Παραδείγματος 9.3 είναι ανεξάρτητες. Παρατηρούμε καταρχήν, πως υπάρχουν ζεύγη τιμών των x, y για τα οποία δεν ισχύει η (9.14) (μπορείτε να βρείτε μερικά;). Άρα, δεν μπορούμε να χρησιμοποιήσουμε το Λήμμα 9.4.

Παρατηρούμε επίσης πως η συνδιακύμανση, που έχουμε υπολογίσει στο Παράδειγμα 9.8, δεν είναι μηδενική, άρα από το άνω λήμμα προκύπτει πως οι X, Y δεν μπορεί να είναι ανεξάρτητες.

Εναλλακτικά, θα μπορούσαμε να χρησιμοποιήσουμε τις περιθώριες πυκνότητες των X, Y, που έχουν βρεθεί στο Παράδειγμα 9.6, ως εξής:

$$\begin{aligned}
P\left(0 \leq X \leq \frac{1}{2}\right) &= \int_0^{\frac{1}{2}} \frac{2}{3}(x+1)\,dx = \frac{2}{3}\left[\frac{x^2}{2}+x\right]_0^{\frac{1}{2}} = \frac{2}{3} \times \left[\frac{1}{8}+\frac{1}{2}\right] = \frac{5}{12}, \\
P\left(0 \leq Y \leq \frac{1}{2}\right) &= \int_0^{\frac{1}{2}} \frac{1}{3}(4y+1)\,dx = \frac{1}{3}\left[2y^2+y\right]_0^{\frac{1}{2}} = \frac{1}{3} \times \left[\frac{1}{2}+\frac{1}{2}\right] = \frac{1}{3}.
\end{aligned}$$

Επομένως,

$$P\left(0 \leq X \leq \frac{1}{2}\right) P\left(0 \leq Y \leq \frac{1}{2}\right) = \frac{5}{36} \neq \frac{1}{8} = P\left(0 \leq X \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right).$$

Η τιμή της πιθανότητας $P\left(0 \leq X \leq \frac{1}{2}, 0 \leq X \leq \frac{1}{2}\right)$ έχει υπολογιστεί στο Παράδειγμα 9.3. Άρα, από τον ορισμό της ανεξαρτησίας προκύπτει ότι οι X, Y δεν είναι ανεξάρτητες, αφού δεν ισχύει η (9.13) για $A = B = \left[0, \frac{1}{2}\right]$.

Παράδειγμα 9.10. Οι X, Y του Παραδείγματος 9.4, των οποίων οι περιθώριες πυκνότητες έχουν υπολογιστεί στο Παράδειγμα 9.5, είναι ανεξάρτητες, αφού ικανοποιούν την (9.14).

Λήμμα 9.5. (Ιδιότητες ανεξάρτητων συνεχών Τ.Μ.) *Έστω X, Y ανεξάρτητες συνεχείς Τ.Μ.*

1. *Έστω συναρτήσεις $g : \mathbb{R} \to \mathbb{R}$, και $h : \mathbb{R} \to \mathbb{R}$. Θα ισχύει*

$$E(g(X)h(Y)) = E(g(X))E(h(Y))).$$

Ειδική περίπτωση της άνω είναι η

$$E(XY) = E(X)E(Y).$$

2. $\mathrm{COV}(X, Y) = 0$.

3. $\mathrm{VAR}(X + Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y)$.

Απόδειξη. 1. Εφαρμόζοντας το Λήμμα 9.2 έχουμε:

$$\begin{aligned} E(g(X)h(Y)) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_{XY}(x,y)\,dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)\,dxdy \\ &= \left(\int_{-\infty}^{\infty} g(x)f_X(x)\,dx\right)\left(\int_{-\infty}^{\infty} h(y)f_Y(y)\,dy\right). \end{aligned}$$

Η δεύτερη ισότητα προκύπτει από την υπόθεση της ανεξαρτησίας. Η τρίτη βγάζοντας σταθερές έξω από ολοκληρώματα.

2. Προκύπτει άμεσα από το προηγούμενα σκέλος και την εφαρμογή της (9.11).

3. Προκύπτει άμεσα από το προηγούμενο σκέλος και την εφαρμογή της (9.12).

□

Ορισμός 9.5. (Ασυσχέτιστες Τ.Μ.) *Δύο συνεχείς Τ.Μ. X, Y, είναι ασυσχέτιστες αν η συνδιακύμανσή τους* $\mathrm{COV}(X, Y)$ *είναι μηδενική, δηλαδή* $\mathrm{COV}(X, Y) = 0$.

Παρατήρηση: Μπορεί να δειχτεί ότι δύο ασυσχέτιστες Τ.Μ. μπορεί να μην είναι ανεξάρτητες, όπως και στη διακριτή περίπτωση.

9.5 Παραδείγματα

Παράδειγμα 9.11. Ο Σταύρος και ο Γιάννης έχουν δώσει ραντεβού σε ένα μπαρ, για τις 03:00 π.μ. Έστω X, Y οι χρόνοι καθυστέρησής τους, σε ώρες. Υποθέτουμε ότι οι Τ.Μ. X, Y είναι ανεξάρτητες η μια από την άλλη, και επιπλέον είναι ομοιόμορφα κατανεμημένες από το 0 ως το 1. Θα απαντήσουμε τα ακόλουθα ερωτήματα:

1. Ποια είναι η πιθανότητα η διαφορά στους χρόνους αφίξεων σε απόλυτη τιμή, δηλαδή το $|X - Y|$, να είναι κάτω από 0.5 ώρες;
2. Πόση είναι η μέση τιμή της διαφοράς των χρόνων άφιξης σε απόλυτη τιμή; Με άλλα λόγια, κατά μέσο όρο πόσο θα περιμένει ο πρώτος που θα έρθει τον δεύτερο;
3. Για να τους αποτρέπει να αργούν, ο μπάρμαν τους χρεώνει στο λογαριασμό, εκτός από τα ποτά, και ένα κόστος $20X + 10Y$ Ευρώ. Κατά μέσο όρο, πόσο πληρώνουν κάθε φορά στον μπάρμαν λόγω της καθυστέρησής τους;
4. Ποια είναι η πιθανότητα να πληρώσουν πάνω από 20 Ευρώ κόστος καθυστέρησης;
5. Πόση είναι η συνδιακύμανση $\mathrm{COV}(X, Y)$;

Για να απαντήσουμε τα ερωτήματα, παρατηρούμε καταρχήν πως οι περιθώριες κατανομές είναι οι

$$f_X(x) = \begin{cases} 1, & x \in [0,1], \\ 0, & x \notin [0,1], \end{cases} \quad f_Y(y) = \begin{cases} 1, & y \in [0,1], \\ 0, & y \notin [0,1], \end{cases}$$

και χρησιμοποιώντας την υπόθεση της ανεξαρτησίας μπορούμε να βρούμε την από κοινού πυκνότητα:

$$f_{XY}(x,y) = f_X(x) f_Y(y) = \begin{cases} 1, & (x,y) \in R = [0,1] \times [0,1], \\ 0, & (x,y) \notin [0,1] \times [0,1]. \end{cases}$$

Έχοντας την από κοινού κατανομή, μπορούμε να απαντήσουμε τα άνω ερωτήματα:

1. $$P\left(|X - Y| < \frac{1}{2}\right) = \iint_{|x-y|<1/2} f_{XY}(x,y)\, dA = \iint_{|x-y|<1/2,\ 0 \le x,y \le 1} 1\, dA.$$

   Στη δεύτερη ισότητα λάβαμε υπ΄ όψιν ότι η ολοκληρωτέα συνάρτηση είναι μηδέν εκτός του τετραγώνου $[0,1] \times [0,1]$, και ίση με τη μονάδα εντός. Το χωρίο $A = \{(x,y) : |x - y| < \frac{1}{2}, 0 \le x, y \le 1\}$ έχει σχεδιαστεί στο Σχήμα 9.12. Επειδή η ολοκληρωτέα συνάρτηση ισούται με τη μονάδα, το ολοκλήρωμα ισούται με το εμβαδόν του A, που είναι $\frac{3}{4}$. Άρα, με πιθανότητα $\frac{3}{4}$ η αναμονή του πρώτου για την άφιξη του δεύτερου δεν θα ξεπεράσει την μισή ώρα.

2. Κατά τα γνωστά από τη θεωρία,

$$\begin{aligned} E(|X-Y|) &= \iint_{\mathbb{R}\times\mathbb{R}} |x-y| f_{XY}(x,y)\, dA = \iint_{[0,1]\times[0,1]} |x-y|\, dA \\ &= \iint_{R_1} (x-y)\, dA + \iint_{R_2} (y-x)\, dA, \end{aligned}$$

όπου

$$R_1 = \{0 \le x \le 1,\ 0 \le y \le x\}, \quad R_2 = \{0 \le x \le 1,\ x \le y \le 1\}.$$

το πρώτο εκ των άνω ολοκληρωμάτων υπολογίζεται ως εξής:

$$\begin{aligned} \iint_{R_1} (x-y)\, dA &= \int_0^1 \left(\int_0^x (x-y)\, dy \right) dx = \int_0^1 \left[xy - \frac{y^2}{2} \right]_0^x dx \\ &= \int_0^1 \frac{x^2}{2}\, dx = \left[\frac{x^3}{6} \right]_0^1 = \frac{1}{6}. \end{aligned}$$

Το δεύτερο ολοκλήρωμα μπορεί να υπολογιστεί παρόμοια, ή εναλλακτικά παρατηρούμε πως λόγω συμμετρίας πρέπει αν ισούται με το πρώτο. Τελικά,

$$E(|X-Y|) = \frac{1}{3}.$$

3. Η μέση τιμή του κόστους είναι εύκολο να υπολογιστεί:

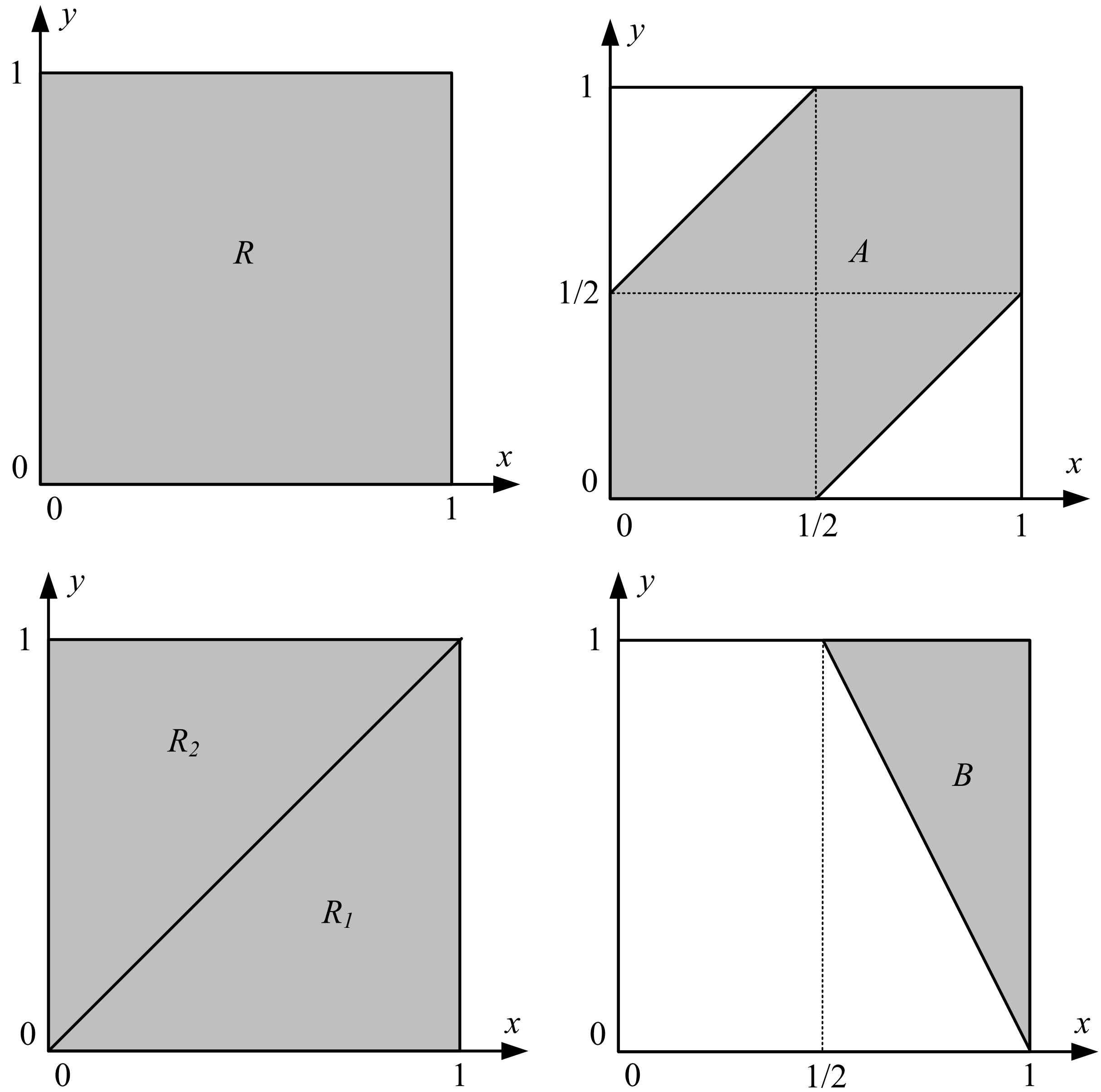
$$E(Z) = E(20X + 10Y) = 20E(X) + 10E(Y) = 20 \times \frac{1}{2} + 10 \times \frac{1}{2} = 15.$$

4. Υπολογίζουμε την πιθανότητα $P(Z > 20)$ ως εξής:

$$\begin{aligned} P(Z > 20) &= P(20X + 10Y > 20) = P(2X + Y > 2) \\ &= \iint_{2x+y>2} f_{XY}(x,y)\, dA = \iint_{2x+y>2,\ 0\le x,y\le 1} 1\, dA. \end{aligned}$$

Και το χωρίο $B = \{(x,y) : 2x + y > 2,\ 0 \le x, y \le 1\}$ εμφανίζεται σκιασμένο στο Σχήμα 9.12. Το εμβαδόν του είναι $\frac{1}{4}$, άρα, επειδή η πυκνότητα είναι σταθερή και ίση με τη μονάδα, έχουμε $P(Z > 20) = \frac{1}{4}$.

5. Αφού οι X, Y είναι ανεξάρτητες, η συνδιακύμανση $\mathrm{COV}(X,Y) = 0$.



Σχήμα 9.12: Τα χωρία του Παραδείγματος 9.11. Η από κοινού πυκνότητα είναι μονάδα εντός του τετραγώνου $R = \{(x, y) : \ 0 \leq x, y \leq 1\}$, και μηδενική εκτός.

Παράδειγμα 9.12. Έστω τώρα πως ο Σταύρος και ο Γιάννης φτάνουν σε τυχαίους χρόνους X και Y αντίστοιχα, για τους οποίους γνωρίζουμε ότι η από κοινού πυκνότητα πιθανότητας είναι η

$$f_{XY}(x,y) = \begin{cases} cx(1-y), & (x,y) \in [0,1]\times[0,1], \\ 0, & (x,y) \notin [0,1]\times[0,1]. \end{cases}$$

Θα απαντήσουμε τα ακόλουθα ερωτήματα:

1. Ποια είναι η τιμή της σταθεράς c;
2. Ποιες είναι οι περιθώριες πυκνότητες των X, Y;
3. Είναι οι Τ.Μ. X, Y ανεξάρτητες;
4. Πόση είναι η συνδιακύμανση $\mathrm{COV}(X,Y)$;
5. Πόση είναι η πιθανότητα $P(X \le Y)$;

Έχουμε:

1. Θα χρησιμοποιήσουμε το γεγονός ότι το ολοκλήρωμα της πυκνότητας σε όλο το επίπεδο πρέπει να ισούται με τη μονάδα. Έχουμε

$$\begin{aligned} \iint\limits_{\mathbb{R}\times\mathbb{R}} f_{XY}(x,y)\,dA &= c\int_0^1 \left(\int_0^1 x(1-y)\,dx\right) dy \\ &= c\int_0^1 (1-y)\left(\int_0^1 x\,dx\right) dy \\ &= c\left(\int_0^1 x\,dx\right)\left(\int_0^1 (1-y)\,dy\right) = c \times \frac{1}{2} \times \frac{1}{2}, \end{aligned}$$

   συνεπώς πρέπει $c = 4$. Η πυκνότητα έχει σχεδιαστεί στο Σχήμα 9.13.

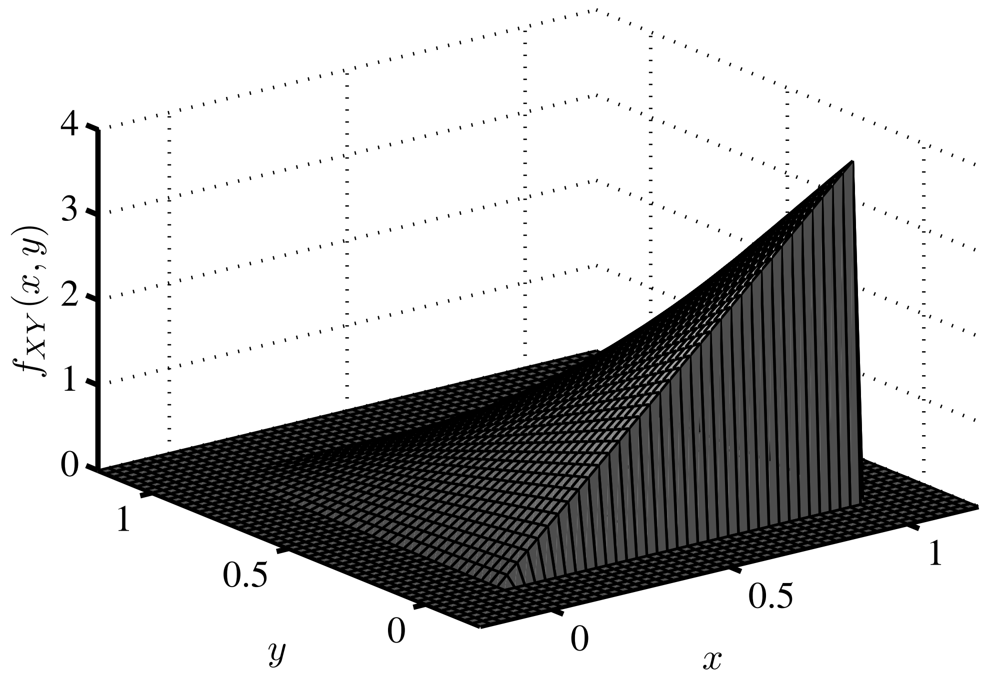
2. Κατά τα γνωστά από τη θεωρία,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy.$$

   Προφανώς όταν $x > 1$ ή $x < 0$, το άνω ολοκλήρωμα είναι μηδέν, γιατί τότε η ολοκληρωτέα συνάρτηση είναι παντού μηδενική.

   Στην περίπτωση που $0 \le x \le 1$, έχουμε:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_0^1 4x(1-y)\,dy = 4x\int_0^1 (1-y)\,dy = 2x.$$



Σχήμα 9.13: Η πυκνότητα πιθανότητας $f_{XY}(x,y)$ του Παραδείγματος 9.12.

Άρα, συγκεντρωτικά

$$f_X(x) = \begin{cases} 2x, & x \in [0,1], \\ 0, & x \notin [0,1]. \end{cases}$$

Παρομοίως, η περιθώρια $f_Y(y) = 0$ όταν $y \notin [0,1]$, ενώ όταν $y \in [0,1]$ έχουμε:

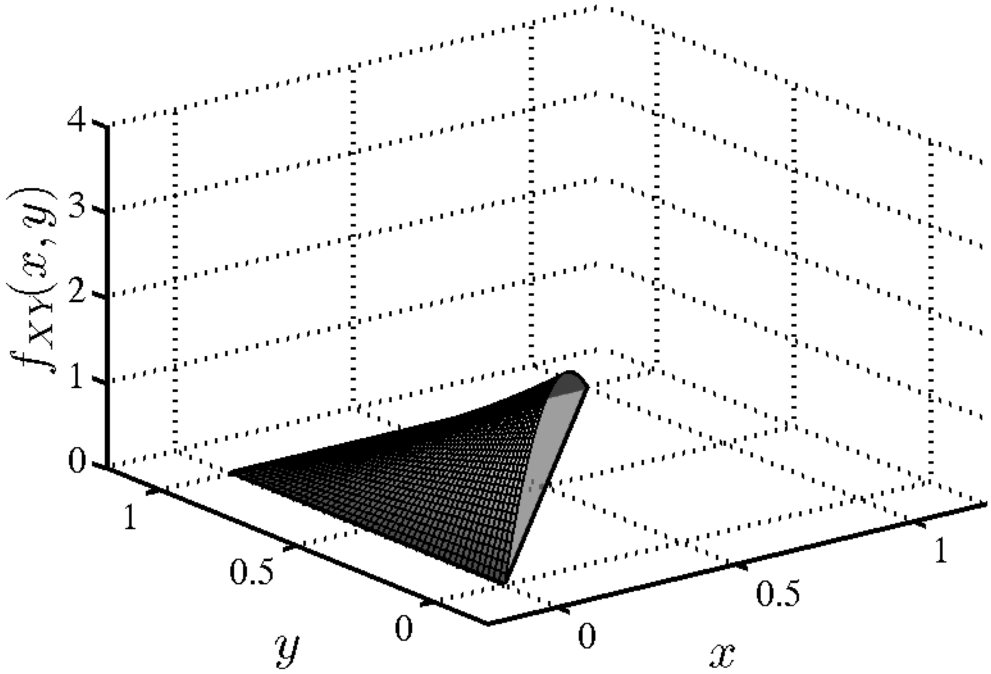
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_0^1 4x(1-y)\,dx = (1-y)\int_0^1 4x\,dx = 2(1-y),$$

και συγκεντρωτικά $f_Y(y) = \begin{cases} 2(1-y), & y \in [0,1], \\ 0, & y \notin [0,1]. \end{cases}$

Για να κατανοήσετε καλύτερα τον υπολογισμό των περιθωρίων πυκνοτήτων, μπορείτε να δείτε το Σχήμα 9.11, που εφαρμόζεται ως έχει και σε αυτό το παράδειγμα.

3. Για κάθε ζεύγος $(x,y) \in \mathbb{R} \times \mathbb{R}$, ισχύει ότι $f_{XY}(x,y) = f_X(x)f_Y(y)$. (Εξετάστε τις συνολικά 9 περιπτώσεις για να σιγουρευτείτε.) Άρα οι X, Y είναι ανεξάρτητες.

4. Αφού οι X, Y είναι ανεξάρτητες, η συνδιακύμανση $\mathrm{COV}(X,Y) = 0$.



Σχήμα 9.14: Ο όγκος του στερεού ισούται με την $P[X \leq Y]$ του Παραδείγματος 9.12.

5.

$$\begin{aligned} P(X \leq Y) &= \iint\limits_{(x,y):x\leq y} f_{XY}(x,y)\,dA = \int_0^1 \left(\int_0^y 4x(1-y)\,dx \right) dy \\ &= 4\int_0^1 \left((1-y)\int_0^y x\,dx \right) dy = 4\int_0^1 (1-y)\frac{y^2}{2}\,dy \\ &= 4\int_0^1 \left(\frac{y^3}{6} - \frac{y^4}{8} \right)' dy = 4\left(\frac{1}{6} - \frac{1}{8} \right) = \frac{1}{6}. \end{aligned}$$

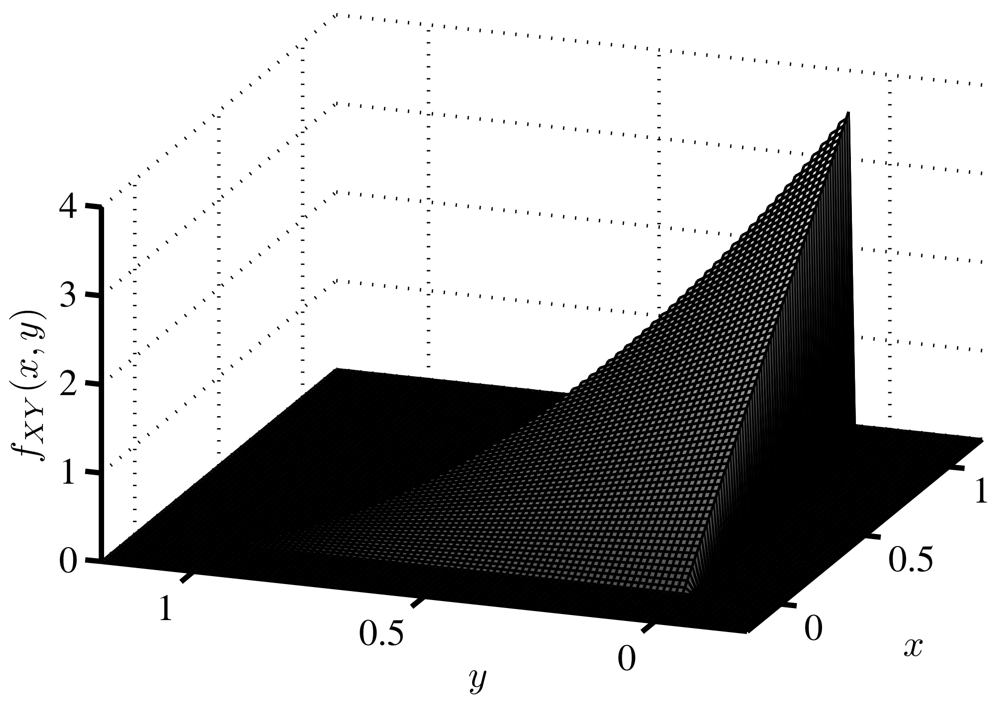
Στο Σχήμα 9.14 έχουμε σχεδιάσει το στερεό του οποίου ο όγκος αντιστοιχεί με την άνω πιθανότητα.

Παράδειγμα 9.13. Έστω, τέλος, πως ο Σταύρος και ο Γιάννης φτάνουν σε τυχαίους χρόνους X και Y αντίστοιχα, για τους οποίους γνωρίζουμε ότι

$$f_{XY}(x,y) = \begin{cases} cx(1-y), & 0 \leq x, y \leq 1,\ x+y \leq 1, \\ 0, & \text{αλλού.} \end{cases}$$

1. Ποια είναι η τιμή της σταθεράς c;

2. Ποιες είναι οι περιθώριες πυκνότητες των X, Y;



Σχήμα 9.15: Η πυκνότητα πιθανότητας του Παραδείγματος 9.13.

3. Είναι οι Τ.Μ. X, Y ανεξάρτητες;
4. Πόση είναι η συνδιακύμανση $\mathrm{COV}(X,Y)$;

Έχουμε, κατά περίπτωση:

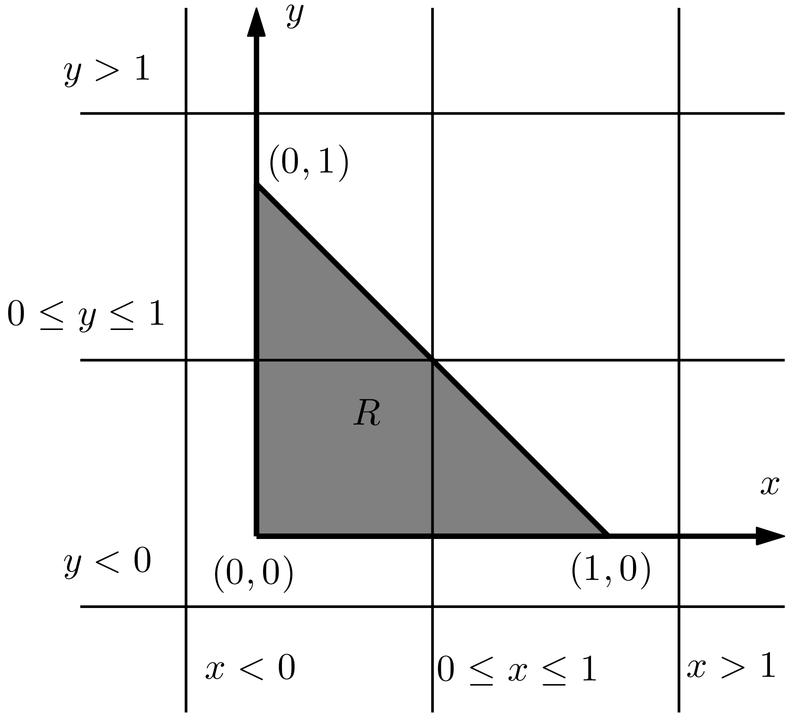
1. Και πάλι, πρέπει το ολοκλήρωμα της $f_{XY}(x,y)$ στο $\mathbb{R}\times\mathbb{R}$ να είναι μονάδα. Όμως,

$$\begin{aligned}
&\iint\limits_{\mathbb{R}\times\mathbb{R}} f_{XY}(x,y)\,dA \\
&= \int_0^1 \left(\int_0^{1-y} cx(1-y)\,dx\right) dy = \int_0^1 \left(c(1-y)\int_0^{1-y} x\,dx\right) dy \\
&= \int_0^1 c(1-y)\frac{(1-y)^2}{2}\,dy = c\int_0^1 \left(-\frac{(1-y)^4}{8}\right)' dy = \frac{c}{8}.
\end{aligned}$$

Άρα τελικά $c = 8$. Η πυκνότητα έχει σχεδιαστεί στο Σχήμα 9.15.

2. Όταν $x \notin [0,1]$, έχουμε $f_X(x) = 0$. Αν $x \in [0,1]$,

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_0^{1-x} 8x(1-y)\,dy \\
&= 8x\int_0^{1-x}(1-y)\,dy = -8x\int_0^{1-x}\left(\frac{(1-y)^2}{2}\right)' dy = 4x(1-x^2),
\end{aligned}$$



Σχήμα 9.16: Παράδειγμα 9.13: Η από κοινού πυκνότητα είναι μη μηδενική στο χωρίο R. Για να υπολογίσουμε τις περιθώριες πυκνότητες $f_X(x)$, $f_Y(y)$, πρέπει να ολοκληρώσουμε την από κοινού κατά μήκος ευθειών καθέτων ή οριζοντίων (αντίστοιχα), έχοντας την μεταβλητή που δεν ολοκληρώνεται, x ή y αντίστοιχα, ως μια παράμετρο που θα επηρεάσει την τιμή του ολοκληρώματος μέσω της θέσης της ευθείας.

και συγκεντρωτικά

$$f_X(x) = \begin{cases} 4x(1-x^2), & x \in [0,1], \\ 0, & x \notin [0,1]. \end{cases}$$

Παρόμοια, αν $y \notin [0,1]$, τότε $f_Y(y) = 0$. Αν $y \in [0,1]$, τότε

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_0^{1-y} 8x(1-y)\,dx \\ &= 4(1-y)\int_0^{1-y} (x^2)'\,dx = 4(1-y)^3. \end{aligned}$$

Τελικά,

$$f_Y(y) = \begin{cases} 4(1-y)^3, & y \in [0,1], \\ 0, & y \notin [0,1]. \end{cases}$$

Στο Σχήμα 9.16 έχουμε σχεδιάσει το χωρίο R στο οποίο η $f_{XY}(x,y)$ είναι μη μηδενική, καθώς και τις διάφορες ευθείες πάνω στις οποίες υπολογίσαμε το (α-

πλό) ολοκλήρωμα της από κοινού πυκνότητας, προκειμένου να υπολογίσουμε τις περιθώριες πυκνότητες.

3. Εύκολα διαπιστώνουμε ότι υπάρχουν ζεύγη (x, y) για τα οποία

$$f_{XY}(x,y) \neq f_X(x) f_Y(y).$$

Παράδειγμα ενός τέτοιου ζεύγους τιμών είναι οι $x = y = \frac{3}{4}$, για το οποίο το αριστερό σκέλος είναι 0, ενώ το δεξί θετικό. Άρα το κριτήριο του Λήμματος 9.4 δεν μπορεί να χρησιμοποιηθεί.

Μπορούμε μάλιστα να διαπιστώσουμε πως οι X, Y δεν είναι ανεξάρτητες παρατηρώντας ότι, ενώ προφανώς

$$P\left(\frac{1}{2} \leq X \leq 1\right) > 0, \qquad P\left(\frac{1}{2} \leq Y \leq 1\right) > 0,$$

έχουμε

$$P\left(\frac{1}{2} \leq X \leq 1, \quad \frac{1}{2} \leq Y \leq 1\right) = \iint_{\frac{1}{2} \leq x,y \leq 1} 0\, dA = 0,$$

άρα η συνθήκη (9.13) δεν ικανοποιείται για $A = B = \left[\frac{1}{2}, 1\right]$.

4. Αφού οι X, Y, δεν είναι ανεξάρτητες, η συνδιακύμανση ενδεχομένως να μην είναι μηδενική, και πρέπει να υπολογιστεί. Έχουμε:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_0^1 4x^2(1-x^2)\, dx = 4\int_0^1 \left(\frac{1}{3}x^3 - \frac{1}{5}x^5\right)' dx \\
&= 4\left[\frac{1}{3}x^3 - \frac{1}{5}x^5\right]_0^1 = \frac{8}{15}, \\
E(Y) &= \int_0^1 y f_Y(y)\, dy = 4\int_0^1 y(1-y)^3\, dy \\
&= 4\int_0^1 (y - 3y^2 + 3y^3 - y^4)\, dy = 4\int_0^1 \left(\frac{1}{2}y^2 - y^3 + \frac{3}{4}y^4 - \frac{1}{5}y^5\right)' dy \\
&= 4\left[\frac{1}{2}y^2 - y^3 + \frac{3}{4}y^4 - \frac{1}{5}y^5\right]_0^1 = \frac{1}{5},
\end{aligned}$$

$$\begin{aligned}
E(XY) &= \iint_{\mathbb{R}\times\mathbb{R}} xy f_{XY}(x,y)\,dA = \iint_{0\le x,y\le 1,\ x+y\le 1} 8x^2(1-y)y\,dA \\
&= \int_0^1 \left(\int_0^{1-x} 8x^2(1-y)y\,dy\right) dx = \int_0^1 8x^2 \left(\int_0^{1-x} (y-y^2)\,dy\right) dx \\
&= \int_0^1 8x^2 \left(\int_0^{1-x} \left(\frac{y^2}{2}-\frac{y^3}{3}\right)' dy\right) dx = \int_0^1 8x^2 \left[\frac{y^2}{2}-\frac{y^3}{3}\right]_0^{1-x} dx \\
&= \int_0^1 8x^2 \left(\frac{1}{2}(1-x)^2 - \frac{1}{3}(1-x)^3\right) dx \\
&= \int_0^1 \frac{4}{3}x^2\left(2x^3-3x^2+1\right) dx = \int_0^1 \left(\frac{8}{3}x^5 - 4x^4 + \frac{4}{3}x^2\right) dx \\
&= \int_0^1 \left(\frac{4}{9}x^6 - \frac{4}{5}x^5 + \frac{4}{9}x^3\right)' dx = \left[\frac{4}{9}x^6 - \frac{4}{5}x^5 + \frac{4}{9}x^3\right]_0^1 = \frac{4}{45},
\end{aligned}$$

και τελικά

$$\mathrm{COV}(X,Y) = E(XY) - E(X)E(Y) = \frac{4}{45} - \frac{8}{15}\times\frac{1}{5} = -\frac{4}{225}.$$

Παράδειγμα 9.14. *(Συνέλιξη)* Έστω δύο ανεξάρτητες συνεχείς Τ.Μ. X, Y με πυκνότητες $f_X(x)$, $f_Y(y)$ και με από κοινού πυκνότητα (δεδομένης της ανεξαρτησίας τους) την $f_{XY}(x,y) = f_X(x)f_Y(y)$. Η πυκνότητα του αθροίσματός τους, έστω $Z = X+Y$, ισούται με

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t)f_Y(z-t)\,dt.$$

(Η πιο πάνω έκφραση είναι η συνεχής μορφή της συνέλιξης δύο συναρτήσεων. Την συναντάμε πολύ συχνά στην επεξεργασία σήματος. Για παράδειγμα, ο ήχος που φτάνει στο αυτί σας από το στερεοφωνικό σας είναι η συνέλιξη της κυματομορφής που είναι αποθηκευμένη στο όποιο μέσο αποθήκευσης (CD, κτλ.) και μιας συνάρτησης που ενσωματώνει τις ρυθμίσεις του equalizer σας, τα χαρακτηριστικά του δωματίου που βρίσκεστε, τη συμπεριφορά των αυτιών σας, κ.ο.κ.)

Για να αποδείξουμε την άνω σχέση, θα υπολογίσουμε πρώτα την κατανομή της Z. Έχουμε:

$$\begin{aligned}
F_Z(z) &= P(Z\le z) = P(X+Y\le z) \\
&= \iint_{(x,y):x+y\le z} f_{XY}(x,y)\,dA = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{z-x} f_X(x)f_Y(y)\,dy\right) dx \\
&= \int_{-\infty}^{\infty} f_X(x)\left(\int_{-\infty}^{z-x} f_Y(y)\,dy\right) dx = \int_{-\infty}^{\infty} f_X(x)F_Y(z-x)\,dx.
\end{aligned}$$

Παραγωγίζοντας, προκύπτει τελικά

$$\begin{aligned} f_Z(z) &= F_Z'(z) = \frac{d}{dz}\left(\int_{-\infty}^{\infty} f_X(x)F_Y(z-x)\,dx\right) \\ &= \int_{-\infty}^{\infty} f_X(x)F_Y'(z-x)\,dx = \int_{-\infty}^{\infty} f_X(t)f_Y(z-t)\,dt. \end{aligned}$$

Στην τρίτη ισότητα, αλλάξαμε την σειρά της ολοκλήρωσης και της παραγώγισης (στην συγκεκριμένη περίπτωση ικανοποιούνται οι συνθήκες που επιτρέπουν αυτή την εναλλαγή.) Στην τέταρτη ισότητα αλλάξαμε την μεταβλητή ολοκλήρωσης από x σε t.

Παράδειγμα 9.15. *(Ελάχιστο και μέγιστο)* Έστω οι ανεξάρτητες συνεχείς Τ.Μ. X, Y με κατανομές $F_X(x)$, $F_Y(y)$ και πυκνότητες $f_X(x)$, $f_Y(y)$. Έστω επίσης οι συνεχείς Τ.Μ. $W = \max\{X, Y\}$ και $V = \min\{X, Y\}$. Θα υπολογίσουμε τις κατανομές $F_W(w)$, $F_V(v)$ και τις πυκνότητες $f_W(w)$, $f_V(v)$, των W και V, συναρτήσει των $F_X(x)$, $F_Y(y)$ και $f_X(x)$, $f_Y(y)$.

Για να βρούμε την κατανομή της W, χρησιμοποιώντας το γεγονός ότι οι X, Y είναι ανεξάρτητες, έχουμε

$$\begin{aligned} F_W(w) &= P(W \le w) = P\big(\{X \le w\} \cap \{Y \le w\}\big) \\ &= P(X \le w)P(Y \le w) = F_X(w)F_Y(w). \end{aligned}$$

Παραγωγίζοντας,

$$f_W(w) = f_X(w)F_Y(w) + F_X(w)f_Y(w).$$

Παρομοίως, για να βρούμε την συνάρτηση κατανομής της V έχουμε:

$$\begin{aligned} F_V(v) &= P(V \le v) = 1 - P(V > v) = 1 - P\big(\{X > v\} \cap \{Y > v\}\big) \\ &= 1 - P(X > v)P(Y > v) = 1 - (1 - F_X(v))(1 - F_Y(v)). \end{aligned}$$

Παραγωγίζοντας:

$$f_V(v) = f_X(v)(1 - F_Y(v)) + f_Y(v)(1 - F_X(v)).$$

Όλα τα αποτελέσματα έχουν διαισθητική εξήγηση. Μπορείτε να τη βρείτε;

9.6 Πολλές Συνεχείς Τυχαίες Μεταβλητές

Όπως και στην περίπτωση των διακριτών Τ.Μ., η θεωρία που έχουμε αναπτύξει μέχρι τώρα επεκτείνεται και στην περίπτωση περισσότερων από δύο συνεχών Τ.Μ. Σε αυτή την παράγραφο παρουσιάζουμε συνοπτικά την επέκταση της θεωρίας, και μερικά χρήσιμα αποτελέσματα της. Παραλείπουμε όλες τις αποδείξεις, καθώς είναι ανάλογες με αυτές της περίπτωσης των δύο Τ.Μ.

Ορισμός 9.6. *Ένα πλήθος n Τ.Μ. $X_1, \ldots, X_n$ καλούνται από κοινού συνεχείς όταν υπάρχει μια συνάρτηση $f_{X_1 \ldots X_n}(x_1, \ldots, x_n) : \mathbb{R}^n \to [0, \infty)$, την οποία καλούμε από κοινού συνάρτηση πυκνότητας πιθανότητας, ή από κοινού πυκνότητα, τέτοια ώστε για οποιοδήποτε χωρίο $R \subseteq \mathbb{R}^n$, να ισχύει ότι*

$$P[(X_1, \ldots, X_n) \in R] = \int \cdots \int_R f_{X_1 \ldots X_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n. \tag{9.15}$$

Λήμμα 9.6. (Ιδιότητες από κοινού πυκνότητας) *Έστω n από κοινού συνεχείς Τ.Μ. $X_1, \ldots, X_n$. Η από κοινού πυκνότητά τους $f_{X_1 \ldots X_n}(x_1, \ldots, x_n)$ έχει τις εξής ιδιότητες:*

1. $\displaystyle \int \cdots \int_{\mathbb{R}^n} f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = 1.$

2. *Η Τ.Μ. X_i είναι συνεχής και έχει περιθώρια πυκνότητα την*

$$f_{X_i}(x_i) = \int \cdots \int_{(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1}} f_{X_1 \ldots X_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n.$$

Λήμμα 9.7. (Ιδιότητες μέσης τιμής συνάρτησης πολλών Τ.Μ.) *Έστω n συνεχείς Τ.Μ. $X_1, \ldots, X_n$ με από κοινού πυκνότητα $f_{X_1 \ldots X_n}(x_1, \ldots, x_n)$.*

1. *Έστω Τ.Μ. $Z = g(X_1, \ldots, X_n)$. Η μέση τιμή της ισούται με*

$$E(Z) = \int \cdots \int_{\mathbb{R}^n} g(x_1, \ldots, x_n) f_{X_1 \ldots X_n}(x_1, \ldots, x_n)\, dx_1 \ldots dx_n.$$

2. $\displaystyle E\left(\sum_{i=1}^{n} a_i X_i + b\right) = \sum_{i=1}^{n} a_i E(X_i) + b.$

3. *Πιο γενικά, αν έχουμε K Τ.Μ. $Z_k = g_k(X_1, \ldots, X_n)$, $k = 1, \ldots, K$, τότε*

$$E\left(\sum_{k=1}^{K} g_k(X_1, \ldots, X_n)\right) = \sum_{k=1}^{K} E(g_k(X_1, \ldots, X_n)).$$

4. $\mathrm{VAR}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{VAR}(X_i) + 2 \sum_{1 \le i < j \le n} \mathrm{COV}(X_i, X_j).$

Ορισμός 9.7. *Ένα πλήθος n από κοινού συνεχών Τ.Μ. $X_1, \ldots, X_n$ είναι ανεξάρτητες αν για οποιαδήποτε υποσύνολα $A_i \subseteq S_{X_i}$, $i = 1, \ldots, n$ ισχύει*

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \ldots P(X_n \in A_n).$$

Λήμμα 9.8. (Κριτήριο ανεξαρτησίας από κοινού συνεχών Τ.Μ.) *Έστω n από κοινού συνεχείς Τ.Μ. $X_1, \ldots, X_n$, με από κοινού πυκνότητα $f_{X_1 \ldots X_n}(x_1, \ldots, x_n)$, και περιθώριες πυκνότητες $f_{X_i}(x_i)$, $i = 1, \ldots, n$. Οι $X_1, \ldots, X_n$ είναι ανεξάρτητες αν*

$$f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n) \quad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Λήμμα 9.9. (Ιδιότητες ανεξάρτητων συνεχών Τ.Μ.) *Έστω n συνεχείς ανεξάρτητες Τ.Μ. $X_1, \ldots, X_n$.*

1. *Οποιοδήποτε υποσύνολο από τις $X_1, \ldots, X_n$ είναι επίσης ανεξάρτητες Τ.Μ.*

2. *Έστω συναρτήσεις $g_i : S_{X_i} \to \mathbb{R}$. Θα ισχύει*

$$E(g_1(X_1) \ldots g_n(X_n)) = E(g_1(X_1)) \ldots E(g_n(X_n)).$$

Ειδική περίπτωση της άνω είναι η

$$E(X_1 \ldots X_n) = E(X_1) \ldots E(X_n).$$

3. $\mathrm{VAR}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{VAR}(X_i).$

Παράδειγμα 9.16. *(Μηνύματα)* Η διάρκεια αποστολής ενός SMS σε δευτερόλεπτα έχει ομοιόμορφη κατανομή στο διάστημα $[1, 3]$, και η διάρκεια αποστολής ενός MMS έχει εκθετική κατανομή με μέση τιμή τα 8 δευτερόλεπτα. Στέλνουμε, σε τρεις ανεξάρτητες αποστολές, 2 SMS και ένα MMS. Θα απαντήσουμε στα ακόλουθα:

1. Ποια είναι η μέση τιμή της συνολικής διάρκειας αποστολής;
2. Ποια είναι η πιθανότητα και τα 2 SMS και το MMS να έχουν όλα διάρκεια πάνω από 2 δευτερόλεπτα το καθένα;
3. Ποια είναι η πιθανότητα το MMS να έχει διάρκεια μεγαλύτερη από τη μέση τιμή της συνολικής διάρκειας των 2 SMS;

Για να απαντήσουμε τα άνω, έστω X_1 η διάρκεια αποστολής του πρώτου SMS, X_2 η διάρκεια αποστολής του δεύτερου SMS, Y η διάρκεια αποστολής του MMS, και Z η συνολική διάρκεια αποστολής των δύο SMS και του MMS, οπότε $Z = X_1 + X_2 + Y$.

1. Για τη μέση τιμή της συνολικής διάρκειας αποστολής έχουμε:
$$E(Z) = E(X_1+X_2+Y) = E(X_1)+E(X_2)+E(Y) = \frac{3+1}{2}+\frac{3+1}{2}+8 = 12,$$
αφού οι X_1, X_2 έχουν ομοιόμορφη κατανομή στο διάστημα $[1,3]$ και η Y έχει εκθετική κατανομή με μέση τιμή 8.

2. Εφόσον οι τρεις αποστολές είναι ανεξάρτητες μεταξύ τους έχουμε ότι η πιθανότητα τα 2 SMS και το MMS να έχουν διάρκεια πάνω από 2 δευτερόλεπτα το καθένα είναι:
$$\begin{aligned} & P(X_1 > 2,\ X_2 > 2,\ Y > 2) \\ = & \ P(X_1 > 2)P(X_2 > 2)P(Y > 2) = \left[\int_2^3 \frac{1}{3-1}\,dx\right]^2 \times e^{-2/8} = \frac{e^{-1/4}}{4}. \end{aligned}$$

3. Η πιθανότητα το MMS να έχει διάρκεια μεγαλύτερη από τη μέση τιμή της συνολικής διάρκειας των 2 SMS είναι
$$P(Y > E(X_1+X_2)) = P(Y > 2+2) = 1 - P(Y \le 4) = e^{-4/8} \simeq 0.6065.$$

Παράδειγμα 9.17. *(Ελάχιστο n εκθετικών Τ.Μ.)* Έστω n Τ.Μ. $X_1,\ldots,X_n$, όλες εκθετικά κατανεμημένες, και με μέσες τιμές $\theta_1,\ldots,\theta_n$. Θα υπολογίσουμε την κατανομή του ελαχίστου τους, $Z = \min\{X_1,\ldots,X_n\}$, υποθέτοντας πως είναι ανεξάρτητες.
$$\begin{aligned} F_Z(z) &= P(Z \le z) = 1 - P(Z > z) = 1 - P\left(X_1 > z,\ldots,X_n > z\right) \\ &= 1 - P(X_1 > z)\ldots P(X_n > z) = 1 - \exp\left[-\frac{z}{\theta_1}\right]\ldots\exp\left[-\frac{z}{\theta_n}\right] \\ &= 1 - \exp\left[-z\left(\sum_{i=1}^n \frac{1}{\theta_i}\right)\right]. \end{aligned}$$

Άρα, το ελάχιστο Z είναι επίσης κατανεμημένο εκθετικά, με μέση τιμή $\left[\sum_{i=1}^n \theta_i^{-1}\right]^{-1}$.

Κεφάλαιο 10

Οριακά Θεωρήματα

10.1 Ανισότητα του Markov

Περιληπτικά, η ανισότητα του Markov λέει πως αν μια Τ.Μ. έχει «μικρή» μέση τιμή τότε δεν μπορεί να παίρνει μεγάλες τιμές με μεγάλη πιθανότητα. Για να κατανοήσουμε αυτό τον αναγκαίο περιορισμό, έστω πως ακούμε ένα βιολόγο να ισχυρίζεται πως «το μέσο βάρος του αφρικανικού (όχι του ευρωπαϊκού!) χελιδονιού είναι 100 γραμμάρια, ενώ το 60% των αφρικανικών χελιδονιών έχει βάρος άνω των 200 γραμμαρίων». Προφανώς ο ισχυρισμός δεν μπορεί να ευσταθεί, γιατί ακόμα και αν το 40% των χελιδονιών δεν είχε καθόλου βάρος, και το υπόλοιπο 60% είχε το ελάχιστο επιτρεπτό, δηλαδή 200, τότε πάλι αυτό το υπόλοιπο 60% θα αρκούσε ώστε το μέσο βάρος να είναι τουλάχιστον $200 \times 0.6 = 120$ γραμμάρια, και θα είχαμε άτοπο. Η γενίκευση ακριβώς αυτού του συλλογισμού οδηγεί στην ανισότητα του Markov και την απόδειξή της.

Λήμμα 10.1. (Ανισότητα του Markov) *Έστω μια Τ.Μ. X που παίρνει πάντα τιμές $X \geq 0$, και έχει μέση τιμή $E(X)$. Τότε:*

$$P(X \geq c) \leq \frac{E(X)}{c}, \quad \forall c \in \mathbb{R},\ c > 0.$$

Απόδειξη. Η ανισότητα ισχύει για όλα τα είδη Τ.Μ. Θα δούμε την απόδειξή της σε δύο ειδικές περιπτώσεις: όταν η X είναι διακριτή, και όταν είναι συνεχής.

Έστω καταρχήν πως η X είναι συνεχής με πυκνότητα $f(x)$. Ξεκινώντας από τον ορισμό της μέσης τιμής,

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx = \int_{0}^{\infty} x\, f(x)\, dx = \int_{0}^{c} x\, f(x)\, dx + \int_{c}^{\infty} x\, f(x)\, dx.$$

Στην δεύτερη ισότητα λάβαμε υπόψη ότι $f(x) = 0$ όταν $x < 0$. Εφόσον όλες οι τιμές x στο πρώτο από τα δύο ολοκληρώματα του δεξιού σκέλους είναι μεγαλύτερες ή ίσες του

μηδενός, το ολοκλήρωμα είναι κι αυτό μεγαλύτερο ή ίσο του μηδενός, συνεπώς

$$E(X) \geq \int_c^\infty x\, f(x)\, dx \geq \int_c^\infty c\, f(x)\, dx = c \int_c^\infty f(x)\, dx,$$

όπου, στην δεύτερη ανισότητα, χρησιμοποιήσαμε το γεγονός ότι στο ολοκλήρωμα που απομένει $x \geq c \Rightarrow xf(x) \geq cf(x)$. Παρατηρώντας, τέλος, πως από τον ορισμό της πυκνότητας το τελευταίο ολοκλήρωμα άνω ισούται με $P(X \geq c)$, έχουμε

$$E(X) \geq c\, P(X \geq c),$$

που είναι η ζητούμενη ανισότητα.

Έστω τώρα η περίπτωση που η X είναι διακριτή, με σύνολο τιμών $S \subseteq [0, \infty)$ και μάζα $p(x)$. Ξεκινώντας από τον ορισμό της μέσης τιμής,

$$E(X) = \sum_{x \in S} x\, p(x) = \sum_{x \in S:\ x < c} x\, p(x) + \sum_{x \in S:\ x \geq c} x\, p(x),$$

όπου χωρίσαμε το άθροισμα σε δύο μέρη, αυτό που αντιστοιχεί σε τιμές $x \in S$ μικρότερες του c, και τις τιμές $x \geq c$. Εφόσον όλες οι τιμές $x \in S$ της X είναι μεγαλύτερες ή ίσες του μηδενός, το πρώτο άθροισμα άνω είναι κι αυτό μεγαλύτερο ή ίσο του μηδενός, συνεπώς

$$E(X) \geq \sum_{x \in S:\ x \geq c} x\, p(x) \geq \sum_{x \in S:\ x \geq c} c\, p(x) = c \sum_{x \in S:\ x \geq c} p(x),$$

όπου, για την δεύτερη ανισότητα, χρησιμοποιήσαμε το γεγονός ότι στο άθροισμα που απομένει όλα τα x είναι μεγαλύτερα ή ίσα του c. Παρατηρώντας, τέλος, πως το τελευταίο άθροισμα άνω ισούται με $P(X \geq c)$ έχουμε

$$E(X) \geq c\, P(X \geq c),$$

που είναι η ζητούμενη ανισότητα. □

Παράδειγμα 10.1. *(Αφρικάνικο χελιδόνι)* Ένας βιολόγος ισχυρίζεται το εξής: «Το μέσο βάρος του αφρικανικού χελιδονιού είναι 100 γραμμάρια, ενώ το 60% των αφρικανικών χελιδονιών έχει βάρος άνω των 200 γραμμαρίων». Βάσει της ανισότητας του Markov, ο ισχυρισμός του δεν ευσταθεί.

Πράγματι, αν X είναι το βάρος ενός τυχαίου χελιδονιού, θα πρέπει να έχουμε,

$$P(X \geq 200) \leq \frac{E(X)}{200} = \frac{100}{200} = \frac{1}{2},$$

που δεν συμφωνεί με τον ισχυρισμό του βιολόγου ότι αυτή η πιθανότητα είναι ίση με 60%.

Παράδειγμα 10.2. *(Γεωμετρική κατανομή)* Έστω μια Τ.Μ. $X \sim$ Γεωμ$(1/5)$, οπότε η X έχει μέση τιμή $E(X) = \frac{1}{1/5} = 5$. Από την ανισότητα του Markov έχουμε πως

$$P(X \geq 15) \leq 5/15 = \frac{1}{3}.$$

Στην πραγματικότητα όμως, η πιο πάνω πιθανότητα είναι (χρησιμοποιώντας γνωστό τύπο για την γεωμετρική κατανομή)

$$P(X \geq 15) = P(X > 14) = \Big(1 - \frac{1}{5}\Big)^{14} \simeq 0.044,$$

δηλαδή σημαντικά μικρότερη.

Παρατήρηση: Το φράγμα που δίνει η ανισότητα του Markov είναι εξαιρετικά χρήσιμο. Για παράδειγμα, όπως θα δούμε σύντομα, χρησιμεύει στην απόδειξη της ανισότητας του Chebychev και μέσω αυτής στην απόδειξη του Νόμου των Μεγάλων Αριθμών. Αλλά σε πολλές περιπτώσεις, όπως στο άνω παράδειγμα, είναι σχετικά ασθενές. Η ισχύς του έγκειται κυρίως στο γεγονός πως *δεν* χρησιμοποιεί κάποια πληροφορία για την κατανομή της X πέραν της μέσης τιμής της $E(X)$, και συνεπώς ισχύει για *οποιαδήποτε* Τ.Μ. X με μέση τιμή $E(X)$.

Παράδειγμα 10.3. *(Χρόνος εκτέλεσης αλγορίθμου)* Ένας αλγόριθμος έχει ως δεδομένα εισόδου έναν ακέραιο αριθμό Y και μία σειρά από n bits, $X_1, X_2, \ldots, X_n$. Ο χρόνος εκτέλεσης του αλγορίθμου είναι

$$T = Y + 2\sum_{i=1}^{n} X_i + \left(\sum_{i=1}^{n} X_i\right)^2 \qquad \text{δευτερόλεπτα.}$$

Αν υποθέσουμε ότι τα $Y, X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες Τ.Μ. όπου $X_i \sim \text{Bern}(1/4)$ και το Y παίρνει τις τιμές $0, 2, 5$ και 9 με πιθανότητα $1/4$ την κάθε μία:

1. Θα υπολογίσουμε τον μέσο χρόνο εκτέλεσης.
2. Θα δείξουμε ότι η πιθανότητα ο χρόνος εκτέλεσης να ξεπερνά τα n^2 δευτερόλεπτα είναι το πολύ
$$\frac{1}{16} + \frac{11}{16n} + \frac{4}{n^2}.$$
(Συνεπώς, για μεγάλα n αυτή η πιθανότητα είναι το πολύ $\simeq \frac{1}{16} = 6.25\%$.)

Πράγματι:

1. Καταρχάς, αν ορίσουμε
$$Z = \sum_{i=1}^{n} X_i,$$
τότε η $Z \sim$ Διων$(n, 1/4)$ και συνεπώς
$$E(Z) = \frac{n}{4} \quad \text{και} \quad \text{VAR}(Z) = n \times \frac{1}{4} \times \Big(1 - \frac{1}{4}\Big) = \frac{3n}{16}.$$
Επιπλέον,
$$E(Z^2) = \text{VAR}(Z) + [E(Z)]^2 = \frac{3n}{16} + \frac{n^2}{16}.$$
Παρατηρώντας πως ο χρόνος εκτέλεσης T μπορεί να εκφραστεί ως $T = Y + 2Z + Z^2$, βρίσκουμε
$$\begin{aligned} E(T) &= E(Y + 2Z + Z^2) = E(Y) + 2E(Z) + E(Z^2) \\ &= \frac{1}{4} \times 0 + \frac{1}{4} \times 2 + \frac{1}{4} \times 5 + \frac{1}{4} \times 9 + 2 \times \frac{n}{4} + \frac{3n}{16} + \frac{n^2}{16} \\ &= 4 + \frac{11n}{16} + \frac{n^2}{16}. \end{aligned}$$

2. Από το προηγούμενο σκέλος και την ανισότητα του Markov έχουμε
$$P(T \geq n^2) \leq \frac{E(T)}{n^2} = \frac{1}{16} + \frac{11}{16n} + \frac{4}{n^2}.$$

Παράδειγμα 10.4. *(Ανισότητα του Chernoff)* Θα δείξουμε πως για οποιαδήποτε Τ.Μ. X και για κάθε $c, \lambda \in \mathbb{R}$, $\lambda > 0$, ισχύει
$$P(X \geq c) \leq \frac{E(e^{\lambda X})}{e^{\lambda c}},$$
εφόσον η μέση τιμή $E(e^{\lambda X})$ υπάρχει. Η ανισότητα είναι μια από τις πιο απλές μορφές της *ανισότητας του Chernoff*. Παρατηρήστε ότι το φράγμα που δίνει μειώνεται εκθετικά με το c, και για αυτό το λόγο η ανισότητα αυτή είναι πολύ χρήσιμη στη Θεωρία Πιθανοτήτων.

Για να την αποδείξουμε, παρατηρήστε πως η Τ.Μ. $e^{\lambda X}$ προφανώς παίρνει πάντα τιμές μεγαλύτερες ή ίσες του μηδενός, και έχει μέση τιμή $E(e^{\lambda X})$. Άρα μπορούμε να εφαρμόσουμε την ανισότητα του Markov για αυτή την Τ.Μ., έτσι ώστε
$$P(X \geq c) = P(\lambda X \geq \lambda c) = P\Big(e^{\lambda X} \geq e^{\lambda c}\Big) \leq \frac{E(e^{\lambda X})}{e^{\lambda c}}$$
και έτσι προκύπτει η ζητούμενη ανισότητα.

Ας εφαρμόσουμε την ανισότητα του Chernoff στην περίπτωση που $X \sim$ Εκθ(θ). Παρατηρήστε πως

$$E(e^{\lambda X}) = \int_{-\infty}^{\infty} e^{\lambda x}\, f(x)\, dx = \int_0^{\infty} e^{\lambda x}\, \frac{1}{\theta}\, e^{-x/\theta}\, dx = \frac{1}{\theta}\int_0^{\infty} e^{-x(1/\theta-\lambda)}\, dx.$$

Έστω πως $\lambda = \frac{1}{\theta}$. Τότε προφανώς $E(e^{\lambda X}) = \infty$. Έστω πως $\lambda \neq \frac{1}{\theta}$. Τότε έχουμε

$$\begin{aligned} E(e^{\lambda X}) &= \frac{1}{\theta}\int_0^{\infty} e^{-x(1/\theta-\lambda)}\, dx = \frac{1}{\theta(1/\theta-\lambda)}\left[-e^{-x(1/\theta-\lambda)}\right]_0^{\infty} \\ &= \frac{1}{\theta(1/\theta-\lambda)}\left[1-\lim_{x\to\infty} e^{-x(1/\theta-\lambda)}\right]. \end{aligned}$$

Το όριο που εμφανίζεται είναι μηδέν για $\lambda < 1/\theta$, και ∞ για $\lambda > 1/\theta$. Άρα, συγκεντρωτικά,

$$E(e^{\lambda X}) = \begin{cases} \infty, & \lambda \geq 1/\theta, \\ \frac{1}{1-\lambda\theta}, & \lambda < \frac{1}{\theta}. \end{cases}$$

Με αντικατάσταση της μέσης τιμής που βρήκαμε στην ανισότητα του Chernoff, λαμβάνουμε τελικά

$$P(X \geq c) \leq \frac{e^{-\lambda c}}{1-\lambda\theta},$$

εφόσον βέβαια $\lambda < 1/\theta$.

Για παράδειγμα, για $\lambda = 1/2 < 1/\theta = 1$ και $c = 10$, το πιο πάνω φράγμα είναι περίπου ίσο με 0.0135, ενώ το φράγμα που μας δίνει η ανισότητα Markov είναι $1/c = 0.1$. Συνεπώς το φράγμα που αποδείξαμε εδώ είναι σημαντικά μικρότερο, και κατά συνέπεια πολύ ισχυρότερο από αυτό που θα μας έδινε η απευθείας χρήση της ανισότητας του Markov.

10.2 Ανισότητα του Chebychev

Το επόμενο αποτέλεσμα (το οποίο προκύπτει από μια σχετικά απλή εφαρμογή της ανισότητας του Markov) λέει πως αν μια Τ.Μ. έχει «μικρή» διασπορά, τότε δεν μπορεί να παίρνει τιμές μακρυά απ' την μέση τιμή της με μεγάλη πιθανότητα.

Λήμμα 10.2. (Ανισότητα του Chebychev) *Έστω μια Τ.Μ. X με μέση τιμή $E(X)$ και διασπορά* $\mathrm{VAR}(X)$. *Τότε:*

$$P(|X-E(X)|\geq c)\leq \frac{\mathrm{VAR}(X)}{c^2},\quad \forall c\in\mathbb{R},\ c>0.$$

Απόδειξη. Έστω μια νέα Τ.Μ. $Y=(X-\mu)^2$ για την οποία, εξ ορισμού, πάντα έχουμε $Y\geq 0$. Από τον ορισμό της διασποράς, παρατηρούμε πως το Y έχει μέση τιμή $E(Y)=E[(X-\mu)^2]=\mathrm{VAR}(X)$. Εφαρμόζοντας την ανισότητα του Markov για την Τ.Μ. Y βρίσκουμε

$$P(|X-\mu|\geq c)=P\left((X-\mu)^2\geq c^2\right)=P\left(Y\geq c^2\right)\leq\frac{E(Y)}{c^2}=\frac{\mathrm{VAR}(X)}{c^2}.$$

□

Παράδειγμα 10.5. Έστω μια Τ.Μ. X με κατανομή Υπερ$(500,150,15)$. Από τις γνωστές ιδιότητες της υπεργεωμετρικής κατανομής, η X έχει μέση τιμή

$$E(X)=\frac{150\times 15}{500}=4.5$$

και διασπορά

$$\mathrm{VAR}(X)=\frac{150\times 15\times(500-150)\times(500-15)}{500^2\times(500-1)}\simeq 3.06162.$$

Έστω τώρα πως θέλουμε να υπολογίσουμε την πιθανότητα $P(|X-4.5|\geq 3)$, δηλαδή την πιθανότητα το X να ισούται με 0 ή με 1 ή με κάποιον ακέραιο από το 8 ως και το 15. Εφόσον γνωρίζουμε την μάζα $p(x)$ της X, μπορούμε να υπολογίσουμε αυτή την πιθανότητα ως $p(0)+p(1)+p(8)+p(9)+\cdots+p(15)$. Αλλά λόγω της πολυπλοκότητας του τύπου της μάζας $p(x)$, και επιπλέον επειδή απαιτεί τον υπολογισμό παραγοντικών $k!$ για μεγάλα k – πράγμα το οποίο συχνά οδηγεί σε σημαντικά αριθμητικά σφάλματα – είναι πολύ πιο εύκολο να χρησιμοποιήσουμε την ανισότητα Chebychev για να βρούμε ένα *φράγμα* για την ζητούμενη πιθανότητα:

$$P(|X-4.5|\geq 3)=P(|X-E(X)|\geq 3)\leq\frac{\mathrm{VAR}(X)}{3^2}\simeq 0.34018.$$

Με βοήθεια υπολογιστή μπορούμε να βρούμε πως

$$P(|X - 4.5| \geq 3) \simeq 0.0807.$$

Παράδειγμα 10.6. Έστω ένας εξυπηρετητής που δέχεται αιτήματα από χρήστες και αποκρίνεται, μετά από κάποια επεξεργασία του κάθε αιτήματος. Από εμπειρικές μετρήσεις γνωρίζουμε πως ο χρόνος απόκρισης X (σε δευτερόλεπτα) έχει μέση τιμή $E(X)$ και τυπική απόκλιση $\sigma = 6$ δευτερόλεπτα.

Θέλουμε να βρούμε ένα «φράγμα» θ τέτοιο ώστε να μπορούμε να εγγυηθούμε ότι, με πιθανότητα τουλάχιστον 99%, η χρονική απόκριση *δεν* θα απέχει από την μέση τιμή πάνω από $\pm\theta$. Δηλαδή, θέλουμε $P(|X - E(X)| < \theta) \geq 0.99$, ή, ισοδύναμα,

$$P(|X - E(X)| \geq \theta) \leq 0.01. \tag{10.1}$$

Αλλά από την ανισότητα του Chebychev έχουμε πως

$$P(|X - E(X)| \geq \theta) \leq \frac{\sigma^2}{\theta^2} = \frac{6^2}{\theta^2}.$$

Συνεπώς, για να ισχύει το ζητούμενο (10.1), αρκεί να διαλέξουμε το θ έτσι ώστε να έχουμε $36/\theta^2 \leq 0.01$, για το οποίο αρκεί να διαλέξουμε $\theta = 60$ δευτερόλεπτα. Με άλλα λόγια, μπορούμε να εγγυηθούμε πως, με πιθανότητα τουλάχιστον 99%, ο χρόνος απόκρισης δεν θα απέχει από τη μέση τιμή για πάνω από ένα λεπτό.

Παράδειγμα 10.7. *(Απόσταση Τ.Μ. από τη μέση τιμή της)* Για μια Τ.Μ. X με άγνωστη κατανομή και με μέση τιμή μ, κάποιος στατιστικολόγος ισχυρίζεται ότι η πιθανότητα η τιμή της X να απέχει από τη μέση τιμή της κατά παραπάνω από 3μ είναι μικρή. Υποστηρίζει αυτό το συμπέρασμα διότι γνωρίζει πως, στη συγκεκριμένη περίπτωση, η διασπορά της X, σ^2, είναι αρκετά μικρότερη από τη μέση τιμή στο τετράγωνο (δηλαδή από το μ^2). Έχει δίκιο στον ισχυρισμό του ή όχι;

Για να απαντήσουμε, παρατηρούμε πως, αφού έχουμε $\sigma^2 \ll \mu^2$ (δηλαδή η διασπορά της X είναι «πολύ μικρότερη» από τη μέση τιμή στο τετράγωνο), από την ανισότητα του Chebychev προκύπτει πως:

$$P(|X - \mu| \geq 3\mu) \leq \frac{\sigma^2}{(3\mu)^2} \ll \frac{\mu^2}{9\mu^2} = \frac{1}{9}.$$

Επομένως, ο στατιστικολόγος έχει δίκιο.

10.3 Ο Νόμος των Μεγάλων Αριθμών

Σε αυτό το σημείο έχουμε πλέον αναπτύξει αρκετά μαθηματικά εργαλεία ώστε να είμαστε σε θέση να διατυπώσουμε και να αποδείξουμε (σε μια απλή μορφή του) ίσως το πιο θεμελιώδες αποτέλεσμα της Θεωρίας Πιθανοτήτων, τον *Νόμο των Μεγάλων Αριθμών*. Χωρίς να μπούμε ακόμα σε μαθηματικές λεπτομέρειες, ο Νόμος των Μεγάλων Αριθμών λέει το εξής:

> Έστω ένα «μεγάλο» πλήθος N από ανεξάρτητες Τ.Μ. $X_1, \ldots, X_N$, οι οποίες έχουν όλες την ίδια κατανομή, και κατά συνέπεια την ίδια μέση τιμή $E(X_i) = \mu$. Τότε
> $$\frac{1}{N}\sum_{i=1}^{N} X_i \ \simeq\ \mu, \quad \text{με πιθανότητα} \simeq 1.$$

Η αυστηρή μαθηματική διατύπωση είναι η ακόλουθη:

Θεώρημα 10.1. (Νόμος των Μεγάλων Αριθμών (Ν.Μ.Α.)) *Έστω μια ακολουθία από ανεξάρτητες Τ.Μ.* $X_1, X_2, \ldots$ *που έχουν όλες την ίδια κατανομή και κατά συνέπεια την ίδια μέση τιμή* $\mu = E(X_i)$ *και την ίδια διασπορά* $\sigma^2 = \mathrm{VAR}(X_i)$. *Έστω ο εμπειρικός μέσος όρος*

$$\bar{X}_N \triangleq \frac{1}{N}\sum_{i=1}^{N} X_i.$$

Για κάθε $\epsilon > 0$, *έχουμε*

$$P\left(|\bar{X}_N - \mu| < \epsilon\right) \to 1, \quad \text{καθώς το } N \to \infty.$$

Απόδειξη. Η απόδειξη είναι μια απλή εφαρμογή της Ανισότητας του Chebychev. Καταρχήν παρατηρούμε πως η μέση τιμή της Τ.Μ. $\bar{X}_N$ είναι

$$E(\bar{X}_N) = E\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N}\sum_{i=1}^{N} E(X_i) = \frac{1}{N} \times N \times \mu = \mu.$$

Παρομοίως, χρησιμοποιώντας το γεγονός ότι τα X_i είναι ανεξάρτητα, η διασπορά της Τ.Μ. $\bar{X}_N$ είναι

$$\begin{aligned}\mathrm{VAR}(\bar{X}_N) &= \mathrm{VAR}\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\mathrm{VAR}\left(\sum_{i=1}^{N} X_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\mathrm{VAR}(X_i)\\ &= \frac{1}{N^2} \times N \times \sigma^2 = \frac{\sigma^2}{N}.\end{aligned}$$

Τέλος, δεδομένου του ότι $\epsilon > 0$, από την ανισότητα του Chebychev έχουμε

$$P\left(|\bar{X}_N - \mu| \geq \epsilon\right) = P\left(|\bar{X}_N - E(\bar{X}_N)| \geq \epsilon\right) \leq \frac{\mathrm{VAR}(\bar{X}_N)}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2}.$$

Άρα,

$$P\left(|\bar{X}_N - \mu| < \epsilon\right) = 1 - P\left(|\bar{X}_N - \mu| \geq \epsilon\right) \geq 1 - \frac{\sigma^2}{N\epsilon^2},$$

το οποίο προφανώς τείνει στο 1, καθώς το $N \to \infty$. □

Παρατηρήσεις

1. Το αποτέλεσμα ισχύει όποιο και αν είναι το είδος των X_i, αρκεί να υπάρχουν η μέση τιμή και η διασπορά τους.

2. Στην παρούσα του μορφή, ο Ν.Μ.Α. μας λέει ότι η ακολουθία των Τ.Μ. (και όχι ακολουθία αριθμών!) $\bar{X}_N$ τείνει στην σταθερά μ καθώς το $N \to \infty$ με την έννοια ότι η πιθανότητα να απέχει έστω και απειροελάχιστα το $\bar{X}_N$ απ΄ τη μ τείνει στο μηδέν. Δηλαδή, για οποιαδήποτε απόκλιση $\epsilon > 0$, οσοδήποτε μικρή, έχουμε

$$P\left(|\bar{X}_N - \mu| \geq \epsilon\right) \to 0, \quad \text{καθώς το } N \to \infty.$$

Στη γλώσσα των πιθανοτήτων, το είδος αυτής της σύγκλισης ονομάζεται «σύγκλιση κατά πιθανότητα».

3. Η συγκεκριμένη διατύπωση του Ν.Μ.Α. καλείται *«Ασθενής Νόμος των Μεγάλων Αριθμών»*. Υπάρχει και ένας *«Ισχυρός Νόμος των Μεγάλων Αριθμών»*, σύμφωνα με τον οποίο

$$P\left(\lim_{N\to\infty} \bar{X}_N = \mu\right) = 1.$$

Στην γλώσσα των πιθανοτήτων, η ακολουθία $\bar{X}_N$ συγκλίνει στο μ «σχεδόν σίγουρα». Οι δύο νόμοι μοιάζουν αρκετά στην διατύπωση, όμως ο Ισχυρός Ν.Μ.Α. αποδεικνύεται πολύ πιο δύσκολα, δεν απαιτεί την ύπαρξη της διασποράς σ^2, και είναι πολύ πιο χρήσιμος σε θεωρητικά προβλήματα πιθανοτήτων.

4. Όπως φαίνεται και από τα επόμενα παραδείγματα, ο Ν.Μ.Α. μας εξασφαλίζει ότι ισχύουν πολλές ιδιότητες των Πιθανοτήτων που διαισθητικά είναι αναμενόμενες, αλλά δεν τις είχαμε αποδείξει μέχρι τώρα. Κατά μία έννοια, ο Ν.Μ.Α. είναι «απόδειξη» ότι η Θεωρία Πιθανοτήτων «δουλεύει»!

Παράδειγμα 10.8. *(Τι θα πει «πιθανότητα»;)* Έστω πως μας ενδιαφέρει η πιθανότητα $p = P(A)$ του να συμβεί ένα συγκεκριμένο ενδεχόμενο A. Λόγου χάρη, το A θα μπορούσε να αντιστοιχεί στα «η θεραπεία του ασθενή ήταν επιτυχής» ή «το δίκτυο παρουσίασε σφάλμα» ή «η εκτέλεση του αλγορίθμου ολοκληρώθηκε κανονικά» κλπ. Ένας

τρόπος για να δώσουμε ένα διαισθητικό νόημα στην πιθανότητα p είναι να φανταστούμε πως επαναλαμβάνεται το ίδιο ακριβώς πείραμα πάρα πολλές, ανεξάρτητες φορές, έτσι ώστε, τουλάχιστον στο διαισθητικό επίπεδο: «η πιθανότητα $p = P(A)$ είναι το ποσοστό των φορών, μακροπρόθεσμα, που συμβαίνει το A». Ο Ν.Μ.Α. είναι ακριβώς το αποτέλεσμα εκείνο που τεκμηριώνει αυτόν το συλλογισμό μαθηματικά, γιατί δείχνει ότι ο τρόπος που ορίσαμε την πιθανότητα είναι συμβατός με την άνω ερμηνεία της πιθανότητας $P(A)$.

Συγκεκριμένα, το πιο πάνω φανταστικό σενάριο μπορεί να περιγραφεί ως εξής. Έστω πως έχουμε ένα μεγάλο πλήθος, N, ανεξάρτητων επαναλήψεων του ίδιο πειράματος, και ας ορίσουμε, για την κάθε επανάληψη i, μια Τ.Μ. X_i που να ισούται με 1 αν το A συνέβη τη φορά i, αλλιώς να ισούται με 0. Τότε,

$$\text{«ποσοστό των φορών που συνέβη το } A\text{»} = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

και ο Ν.Μ.Α. μας λέει πως, με πιθανότητα που τείνει στη μονάδα, αυτό θα απέχει από την μέση τιμή των X_i λιγότερο από ϵ. Εφόσον τα X_i είναι Bernoulli Τ.Μ., η μέση τιμή $E(X_i)$ είναι απλά η πιθανότητα $p = P(A)$ του να έχουμε $X_i = 1$, δηλαδή του να συμβεί το A. Με άλλα λόγια:

> *Η πιθανότητα p του να συμβεί ένα οποιοδήποτε ενδεχόμενο A, ισούται με τη συχνότητα με την οποία θα συμβεί, μακροπρόθεσμα, το A, σε πολλές ανεξάρτητες επαναλήψεις του ίδιου πειράματος.*

Παράδειγμα 10.9. *(Γιατί γίνονται δημοσκοπήσεις;)* Έστω πως σε έναν πληθυσμό M ατόμων οι $M/4$ είναι ψηφοφόροι κάποιου κόμματος, δηλαδή το κόμμα αυτό στις εκλογές θα πάρει ποσοστό 25%. Μια δημοσκόπηση επιλέγει στην τύχη ένα «μεγάλο» πλήθος ατόμων N (με επανατοποθέτηση) και τα ρωτά αν θα ψηφίσουν το κόμμα αυτό. (Σε τυπικές δημοσκοπήσεις έχουμε N μεταξύ 1000 και 2000 ερωτηθέντων σε έναν πληθυσμό M μεταξύ 500 χιλιάδων και 8 εκατομμυρίων ατόμων.)

Έστω X_i μια Τ.Μ. που ισούται με 1 αν το άτομο i που ρωτάται είναι ψηφοφόρος του κόμματος, και 0 αν όχι, οπότε οι X_i είναι ανεξάρτητες Τ.Μ. Bern(0.25). Τότε,

$$\text{«ποσοστό ερωτηθέντων που είναι ψηφοφόροι του κόμματος»} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

και ο Ν.Μ.Α. μας λέει πως, με πιθανότητα που τείνει στη μονάδα, αυτό θα απέχει από την μέση τιμή των X_i, δηλαδή το 0.25, λιγότερο από ϵ. Άρα, από ένα αρκετά μεγάλο (τυχαίο) δείγμα, μπορούμε να εκτιμήσουμε το πραγματικό ποσοστό ψήφων που αυτό το κόμμα θα πάρει στις εκλογές.

Παράδειγμα 10.10. *(Δειγματοληψία)* Έστω πως θέλουμε να υπολογίσουμε το μέσο ύψος $\bar{y}$ (σε εκατοστά) που έχουν M κορίτσια σε ένα σχολείο. Αν y_k είναι το ύψος του κοριτσιού $k = 1, 2, \ldots, M$, θέλουμε να εκτιμήσουμε το

$$\text{«μέσο ύψος»} = \bar{y} = \frac{1}{M}\sum_{k=1}^{M} y_k.$$

Αντί να εξετάσουμε ολόκληρο τον πληθυσμό, επιλέγουμε τυχαία, με επανατοποθέτηση, N μέλη του πληθυσμού, όπου το μέγεθος N του δείγματός μας είναι μεν σχετικά μεγάλο, αλλά είναι σημαντικά μικρότερου του συνολικού μεγέθους M του πληθυσμού (όπως και στο προηγούμενο παράδειγμα). Έστω X_j το ύψος του κοριτσιού j που επιλέξαμε. Τότε τα X_j είναι ανεξάρτητα, και το κάθε X_j έχει σύνολο τιμών $S = \{y_1, y_2, \ldots, y_M\}$ και μάζα

$$p(y_k) = P(X_j = y_k) = P(\text{«επιλέξαμε το κορίτσι } k\text{»}) = \frac{1}{M}.$$

(Για απλότητα, υποθέσαμε ότι όλα τα ύψη είναι διαφορετικά, ας πούμε γιατί τα έχουμε μετρήσει με άπειρη ακρίβεια.) Άρα τα X_j έχουν μέση τιμή

$$E(X_j) = \sum_{k=1}^{M} y_k \times p(y_k) = \sum_{k=1}^{M} y_k \times \frac{1}{M} = \frac{1}{M}\sum_{k=1}^{M} y_k = \bar{y}.$$

Αν εκτιμήσουμε λοιπόν το $\bar{y}$ ως τον μέσο όρο απ᾽ τα ύψη των κοριτσιών που επιλέξαμε, ο Ν.Μ.Α. διαβεβαιώνει πως, αν το μέγεθος N του δείγματός μας είναι αρκετά μεγάλο, τότε η εκτίμησή μας θα είναι ακριβής με πιθανότητα κοντά στο 100%:

$$\text{«μέσο ύψος δειγμάτων»} = \frac{1}{N}\sum_{i=1}^{N} X_i \simeq E(X_i) = \bar{y}, \quad \text{με πιθανότητα} \simeq 1.$$

Παράδειγμα 10.11. *(Κόστος συντήρησης αυτοκινήτου)* Έστω X_i ο χρόνος (σε έτη) ανάμεσα στην $i-1$ και την i βλάβη ενός αυτοκινήτου. (Έστω πως X_1 είναι ο χρόνος για την πρώτη βλάβη.) Έστω επίσης Y_i το κόστος της βλάβης i, σε ευρώ. Δίνεται ότι οι X_i έχουν όλες την ίδια κατανομή, με μέση τιμή μ_X, και είναι ανεξάρτητες μεταξύ τους. Επιπλέον, οι Y_i έχουν επίσης όλες την ίδια κατανομή, με μέση τιμή μ_Y, και είναι ανεξάρτητες μεταξύ τους. Από τον Νόμο των Μεγάλων Αριθμών, καθώς $n \to \infty$,

$$\frac{\sum_{i=1}^{n} X_i}{n} \to \mu_X, \quad \frac{\sum_{i=1}^{n} Y_i}{n} \to \mu_Y.$$

Παίρνοντας το πηλίκο, έχουμε

$$\frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i} \to \frac{\mu_Y}{\mu_X}.$$

Το αριστερό σκέλος εκφράζει το μέσο κόστος, σε ευρώ ανά έτη, και βρήκαμε ότι η ποσότητα αυτή συγκλίνει στο $\frac{\mu_Y}{\mu_X}$, όπως ήταν διαισθητικά αναμενόμενο.

10.4 Το Κεντρικό Οριακό Θεώρημα

Έστω X_i, $i = 1, 2, \ldots$ μια ακολουθία από ανεξάρτητες Τ.Μ. με την ίδια κατανομή, μέση τιμή $E(X) = \mu$ και διασπορά $\text{VAR}(X) = \sigma^2$.

Το πρώτο θεμελιώδες αποτέλεσμα των πιθανοτήτων, ο Ν.Μ.Α., μας εξασφαλίζει ότι η πιθανότητα ο εμπειρικός μέσος όρος

$$\bar{X}_N \triangleq \frac{1}{N}\sum_{i=1}^{N} X_i$$

να παρουσιάζει απόκλιση από το μ μεγαλύτερη από ϵ τείνει στο 0, δηλαδή

$$\lim_{N\to\infty} P(|\bar{X}_N - \mu| > \epsilon) = 0.$$

Δύο βασικά και μάλλον προφανή ερωτήματα που γεννιούνται από τον Ν.Μ.Α. είναι τα εξής:

1. Πόσο μεγάλο πρέπει να είναι το πλήθος N των δειγμάτων, ώστε να έχουμε κάποια σχετική βεβαιότητα πως ο εμπειρικός μέσος όρος $\bar{X}_N$ θα είναι αρκετά κοντά στη μέση τιμή μ;

2. Επιπλέον, δεδομένου του πλήθους N, πόσο μικρή είναι η πιθανότητα το $\bar{X}_N$ να απέχει κατά πολύ από το μ;

Αν εξετάσουμε προσεκτικά την απόδειξη του Ν.Μ.Α. που δώσαμε θα δούμε πως, πέρα απ' το ασυμπτωτικό αποτέλεσμα του θεωρήματος (το οποίο ισχύει μόνο καθώς το $N \to \infty$), η απόδειξη περιέχει και κάποιες πρώτες απαντήσεις στα πιο πάνω ερωτήματα. Συγκεκριμένα, δείξαμε πως η μέση τιμή και η διασπορά του $\bar{X}_N$ είναι, αντίστοιχα,

$$E(\bar{X}_N) = \mu \quad \text{και} \quad \text{VAR}(\bar{X}_N) = \frac{\sigma^2}{N}, \quad \text{για } \textit{κάθε } N.$$

Επιπλέον βρήκαμε ένα ακριβές ποσοτικό φράγμα για την πιθανότητα ο εμπειρικός μέσος όρος $\bar{X}_N$ να απέχει από την μέση τιμή μ κατά τουλάχιστον ϵ:

$$P\left(|\bar{X}_N - \mu| \geq \epsilon\right) \leq \frac{\sigma^2}{N\epsilon^2}.$$

Αν και μαθηματικά σωστό, το πιο πάνω φράγμα στις περισσότερες περιπτώσεις δεν είναι αρκετά ακριβές ώστε να είναι χρήσιμο στην πράξη. Με άλλα λόγια, η ζητούμενη πιθανότητα στην άνω σχέση είναι συνήθως σημαντικά μικρότερη από το φράγμα $\sigma^2/(N\epsilon^2)$.

Το δεύτερο θεμελιώδες αποτέλεσμα των πιθανοτήτων, το *Κεντρικό Οριακό Θεώρημα* καλύπτει αυτή την αδυναμία του Ν.Μ.Α., δίνοντας μια ακριβή προσέγγιση της πιθανότητας $P\left(|\bar{X}_N - \mu| \geq \epsilon\right)$.

Συγκεκριμένα, το Κεντρικό Οριακό Θεώρημα λέει πως, κάτω από ορισμένες συνθήκες, η κατανομή του εμπειρικού μέσου όρου $\bar{X}_N$ μπορεί να προσεγγιστεί με μεγάλη ακρίβεια από την *κανονική κατανομή* με παραμέτρους μ και $\frac{\sigma^2}{N}$, όπου σ^2 είναι η διασπορά των X_i.

Ισοδύναμα, επειδή για τις κανονικές Τ.Μ. γνωρίζουμε ότι αν η $X \sim N(\mu, \sigma^2)$ τότε η $Y = \frac{X-\mu}{\sigma} \sim N(0,1)$, θα ισχύει επίσης ότι το άθροισμα

$$\frac{\bar{X}_N - \mu}{\sqrt{\sigma^2/N}} = \frac{\sqrt{N}}{\sigma}\left[\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) - \mu\right] = \frac{1}{\sigma\sqrt{N}}\sum_{i=1}^{N}(X_i - \mu)$$

ακολουθεί κατά προσέγγιση την τυπική κανονική κατανομή $N(0,1)$. Συνεπώς, οι πιθανότητες ενδεχόμενων που αφορούν τον $\bar{X}_N$ μπορούν εύκολα να υπολογιστούν.

Θεώρημα 10.2. (Κεντρικό Οριακό Θεώρημα (Κ.Ο.Θ.)) *Έστω μια ακολουθία από ανεξάρτητες Τ.Μ. $X_1, X_2, \ldots$ που έχουν όλες την ίδια κατανομή και κατά συνέπεια την ίδια μέση τιμή $\mu = E(X_i)$ και την ίδια διασπορά $\sigma^2 = \mathrm{VAR}(X_i)$. Έστω το κανονικοποιημένο άθροισμα*

$$\bar{S}_N \triangleq \frac{\left(\frac{1}{N}\sum_{i=1}^{N} X_i\right) - \mu}{\sigma\sqrt{1/N}} = \frac{1}{\sigma\sqrt{N}}\sum_{i=1}^{N}(X_i - \mu) = \frac{\left(\sum_{i=1}^{N} X_i\right) - N\mu}{\sigma\sqrt{N}}.$$

Έστω επίσης $a, b \in \mathbb{R}$, $a < b$. Ισχύουν τα ακόλουθα, καθώς $N \to \infty$:

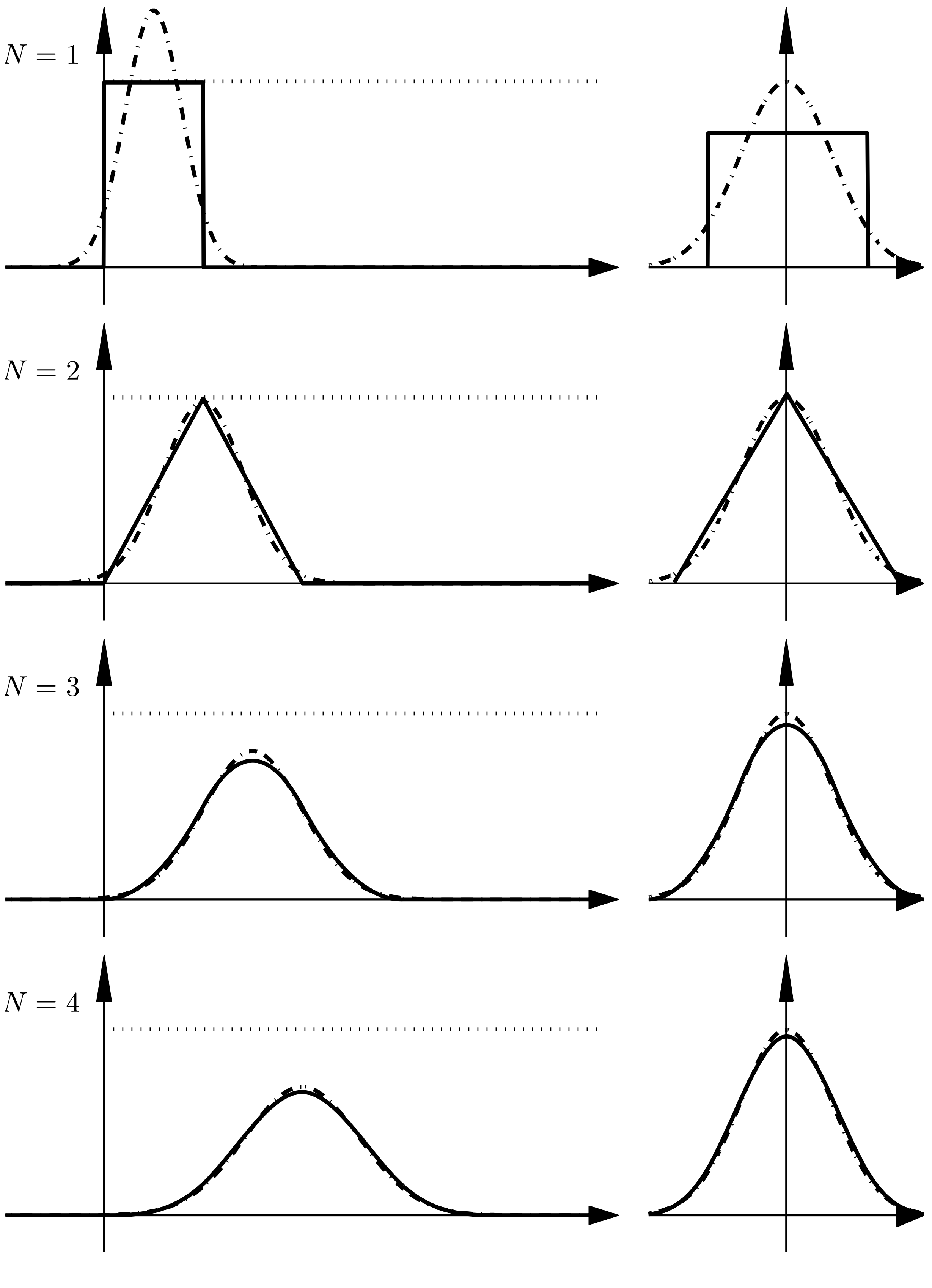
$$\begin{aligned} P\left(a \le \bar{S}_N \le b\right) &\to \Phi(b) - \Phi(a), \\ P\left(\bar{S}_N \le b\right) &\to \Phi(b), \\ P\left(\bar{S}_N \ge a\right) &\to 1 - \Phi(a). \end{aligned}$$

Παρατηρήσεις

1. Η απόδειξη του Κ.Ο.Θ. είναι πολύ τεχνική, και ξεφεύγει κατά πολύ από τους στόχους του παρόντος μαθήματος. Συνεπώς την παραλείπουμε.

2. Το Κ.Ο.Θ. δεν είναι μόνο στο διαισθητικό επίπεδο μια ακριβέστερη μορφή του Ν.Μ.Α., αλλά μπορούμε και να αποδείξουμε τον Ν.Μ.Α. από το Κ.Ο.Θ.

3. Παρατηρήστε ότι έχουμε ένα είδος σύγκλισης που δεν έχουμε ξαναδεί. Συγκεκριμένα, η κατανομή μιας ακολουθίας Τ.Μ. $\bar{S}_N$ συγκλίνει σε μια άλλη κατανομή, στην συγκεκριμένη περίπτωση την τυπική κανονική. Αυτού του είδους η σύγκλιση καλείται «σύγκλιση κατά κατανομή».



Σχήμα 10.1: Σύγκλιση της κατανομής των αθροισμάτων $\sum_{i=1}^{N} X_i$ (αριστερά) και των κανονικοποιημένων αθροισμάτων $\bar{S}_N = \frac{\left(\sum_{i=1}^{N} X_i\right) - N\mu}{\sigma\sqrt{N}}$ (δεξιά) στην κανονική κατανομή, στην περίπτωση $X_i \sim U[0,1]$.

4. Ένα παράδειγμα της σύγκλισης που προβλέπει το Κ.Ο.Θ. φαίνεται στο Σχήμα 10.1. Αριστερά φαίνονται με συνεχή γραμμή οι πυκνότητες, όπως έχουν υπολογιστεί με χρήση Η/Υ, των αθροισμάτων $\sum_{i=1}^{N} X_i$, για $N = 1, 2, 3, 4$, και όταν τα X_i είναι κατανεμημένα ομοιόμορφα στο $[0, 1]$. Φαίνονται επίσης, με διακεκομμένη γραμμή, οι κανονικές κατανομές στις οποίες συγκλίνουν οι πυκνότητες. Στο δεξί μέρος έχουμε σχεδιάσει τις πυκνότητες των κανονικοποιημένων αθροισμάτων $\bar{S}_N = \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^{N}(X_i - \mu)$ και τις αντίστοιχές τους κανονικές. Παρατηρήστε ότι η σύγκλιση είναι πολύ καλή, ακόμα και για την μικρή τιμή $N = 4$.

5. Η πρακτική χρησιμότητα του Κ.Ο.Θ. είναι ότι μας επιτρέπει να υπολογίζουμε προσεγγιστικά πιθανότητες που αφορούν μέσους όρους και/ή αθροίσματα ενός πλήθους N ανεξάρτητων Τ.Μ. $X_1, X_2, \ldots, X_N$, όλων με την ίδια κατανομή. Η μεθοδολογία είναι πάντα η ίδια:

   (α΄) Υπολογίζουμε (αν δεν γνωρίζουμε ήδη) τη μέση τιμή και τη διασπορά των X_i.

   (β΄) Εκφράζουμε το ενδεχόμενο του οποίου την πιθανότητα θέλουμε να υπολογίσουμε ως μια ανισότητα ή συνδυασμό ανισοτήτων που αφορούν το κανονικοποιημένο άθροισμα $\bar{S}_N$, σε οποιαδήποτε από τις μορφές του μας βολεύει.

   (γ΄) Επικαλούμαστε το Κ.Ο.Θ. για να προσεγγίσουμε την πιθανότητα του ενδεχόμενου που βρήκαμε στο προηγούμενο βήμα μέσω της κανονικής κατανομής.

6. Το σφάλμα που κάνουμε στην άνω προσέγγιση εξαρτάται από πολλούς παράγοντες:

   (α΄) Την απόσταση μεταξύ των a, b.

   (β΄) Την σχέση ανάμεσα στα a, b, μ.

   (γ΄) Την κατανομή των X_i.

   (δ΄) Την τιμή του N.

   Γενικώς, όταν το N είναι αρκετά μεγάλο, πχ μεγαλύτερο του 50, και όταν τα a, b απέχουν αρκετά μεταξύ τους και είναι αρκετά κοντά στο μ, ώστε η πιθανότητα που βρίσκουμε να μην είναι πολύ μικρή, π.χ., μεγαλύτερη του 0.01, τότε η προσέγγιση είναι ικανοποιητική.

7. Δείτε τα ακόλουθα παραδείγματα για εφαρμογές αυτής της μεθοδολογίας.

Παράδειγμα 10.12. Έστω πως X_i είναι η θερμοκρασία ενός επεξεργαστή την ημέρα i, όπου θεωρούμε πως τα X_i είναι ανεξάρτητες Τ.Μ. με κατανομή $X_i \sim U[10, 50]$. Ποια η πιθανότητα η μέση θερμοκρασία του επεξεργαστή σε μία περίοδο 3 μηνών (δηλαδή 90 ημερών) να ξεπερνά τους 31 βαθμούς;

Για να εφαρμόσουμε το Κ.Ο.Θ. αρχικά υπολογίζουμε την μέση τιμή και την διασπορά των X_i οι οποίες, από τις αντίστοιχες ιδιότητες της ομοιόμορφης κατανομής είναι

$$\mu = E(X_i) = \frac{10+50}{2} = 30, \quad \sigma^2 = \mathrm{VAR}(X_i) = \frac{(50-10)^2}{12} = \frac{400}{3}$$

αντίστοιχα. Άρα, η ζητούμενη πιθανότητα μπορεί να εκφραστεί ως προς το κανονικοποιημένο άθροισμα του Κ.Ο.Θ. ως εξής:

$$P\Big(\frac{1}{N}\sum_{i=1}^{N} X_i > 31\Big) = P\Big(\frac{1}{N}\sum_{i=1}^{N}(X_i-\mu) > 31-30\Big) = P\Big(\frac{1}{\sigma\sqrt{N}}\sum_{i=1}^{N}(X_i-\mu) > \frac{\sqrt{N}}{\sigma}\Big),$$

όπου η τιμή $\sqrt{N}/\sigma$ είναι $\sqrt{90}/\sqrt{400/3} \simeq 0.8216$. Άρα, από το Κ.Ο.Θ., η ζητούμενη πιθανότητα μπορεί να προσεγγιστεί ως

$$P\Big(\frac{1}{N}\sum_{i=1}^{N} X_i > 31\Big) \simeq P(Z > 0.82) = 1 - P(Z \le 0.82) = 1 - \Phi(0.82),$$

όπου η Z έχει κατανομή $N(0,1)$. Αντικαθιστώντας την τιμή $\Phi(0.82) \simeq 0.7939$, τελικά βρίσκουμε πως η ζητούμενη πιθανότητα είναι $\simeq 1 - 0.7939 = 0.2061$.

Παράδειγμα 10.13. Από εμπειρικές μετρήσεις παρατηρούμε πως η διάρκεια εκτέλεσης ενός αλγορίθμου διαρκεί κατά μέσο όρο 17.5 δευτερόλεπτα, με τυπική απόκλιση ± 4 δευτερόλεπτα. Ποια είναι η πιθανότητα, σε 400 διαδοχικές, ανεξάρτητες χρήσεις του αλγορίθμου, ο συνολικός χρόνος εκτέλεσης να μην ξεπερνά τις δύο ώρες;

Έστω X_i ο χρόνος εκτέλεσης του αλγορίθμου την φορά i για $i = 1, \ldots, N = 400$. Από τα δεδομένα του προβλήματος έχουμε πως τα X_i είναι ανεξάρτητες Τ.Μ. με μέση τιμή $\mu = 17.5$ δευτερόλεπτα και τυπική απόκλιση $\sigma = 4$ δευτερόλεπτα. Συνεπώς, η ζητούμενη πιθανότητα μπορεί να εκφραστεί χρησιμοποιώντας το κανονικοποιημένο άθροισμα του Κ.Ο.Θ. ως εξής:

$$\begin{aligned}
P\Big(\sum_{i=1}^{N} X_i \le 2\times 60\times 60\Big) &= P\Big(\sum_{i=1}^{N}(X_i-\mu) \le 7200 - 400\times 17.5\Big) \\
&= P\Big(\frac{1}{\sigma\sqrt{N}}\sum_{i=1}^{N}(X_i-\mu) \le \frac{200}{4\times\sqrt{400}}\Big) \\
&\simeq P(Z \le 2.5),
\end{aligned}$$

όπου $Z \sim N(0,1)$ και πάλι εφαρμόσαμε το Κ.Ο.Θ. Άρα η ζητούμενη πιθανότητα είναι

$$P\Big(\sum_{i=1}^{N} X_i \le 7200\Big) \simeq P(Z \le 2.5) = \Phi(2.5) \simeq 0.9938,$$

όπου αντικαταστήσαμε την τιμή $\Phi(2.5) \simeq 0.9938$.

Παράδειγμα 10.14. *(Διαγώνισμα)* Σε ένα διαγώνισμα υπάρχουν 150 ερωτήσεις multiple choice, η κάθε μια με 5 δυνατές απαντήσεις. Για κάθε σωστή απάντηση ο εξεταζόμενος παίρνει 3 βαθμούς και για κάθε λάθος χάνει 1 βαθμό. Υποθέτουμε ότι κάποιος απαντάει όλες τις ερωτήσεις στην τύχη, ανεξάρτητα τη μια απ' την άλλη. Ποια η πιθανότητα να πάρει μέσο όρο βαθμολογίας μικρότερο του $\frac{1}{6}$;

Η πιθανότητα να επιλέξει τη σωστή απάντηση είναι 1/5 και η πιθανότητα να επιλέξει λάθος είναι 4/5. Ορίζουμε τις παρακάτω Τ.Μ., για $i = 1, \ldots, 150$:

$$X_i = \begin{cases} 3, & \text{με πιθανότητα } p = 1/5, \\ -1, & \text{με πιθανότητα } 1 - p = 4/5. \end{cases}$$

Έχουμε ότι:

$$E(X_i) = 3 \times \frac{1}{5} - 1 \times \frac{4}{5} = -\frac{1}{5}$$

και

$$\text{VAR}(X_i) = E(X_i^2) - (E(X_i))^2 = \left(3^2 \times \frac{1}{5} + (-1)^2 \times \frac{4}{5}\right) - \left(-\frac{1}{5}\right)^2 = \frac{64}{25}.$$

Η πιθανότητα να πάρει ο εξεταζόμενος τελικά μέσο όρο βαθμολογίας κάτω του $\frac{1}{6}$ είναι

$$\begin{aligned} P\left(\frac{1}{150}\sum_{i=1}^{150} X_i < \frac{1}{6}\right) &= P\left(\sum_{i=1}^{150} X_i < 25\right) = P\left(\sum_{i=1}^{150}(X_i - (-1/5)) < 55\right) \\ &= P\left(\frac{\sum_{i=1}^{150}(X_i - (-1/5))}{\sqrt{150 \times 64/25}} < \frac{55}{\sqrt{150 \times 64/25}}\right) \\ &= \Phi\left(\frac{55}{\sqrt{150 \times 64/25}}\right) \simeq \Phi(2.81) \simeq 0.9975. \end{aligned}$$

Παρατήρηση: Η άνω μεθοδολογία μπορεί να εφαρμοστεί για οποιοδήποτε είδος κατανομής των X_i. Στην ειδική περίπτωση που οι X_i λαμβάνουν συνεχόμενες ακέραιες τιμές, μπορούμε να κάνουμε μια μικρή τροποποίηση, που βελτιώνει κάπως την ακρίβεια της προσέγγισης, ιδιαιτέρως όταν η ζητούμενη πιθανότητα είναι πολύ μικρή. Συγκεκριμένα, έστω $m < M$ και έστω πως πρέπει να υπολογίσουμε την πιθανότητα

$$P\left(m \le \sum_{i=1}^{N} X_i \le M\right).$$

Τότε, προσεγγίζουμε αντί αυτής την πιθανότητα

$$P\left(m - \frac{1}{2} \le \sum_{i=1}^{N} X_i \le M + \frac{1}{2}\right),$$

Παρατηρήστε ότι αφού οι X_i είναι ακέραιες, θα είναι και το άθροισμά τους ακέραιο, και άρα οι δύο πιθανότητες είναι ακριβώς ίσες. Οι προσεγγίσεις τους όμως θα διαφέρουν μεταξύ τους, και η προσέγγιση που θα βασίζεται στο δεύτερο ενδεχόμενο θα είναι καλύτερη. Διαισθητικά, είναι λογικό να αντιστοιχίσουμε στο ενδεχόμενο $\sum_{i=1}^{N} X_i = m$ το διάστημα $(m-\frac{1}{2}, m+\frac{1}{2})$, στο ενδεχόμενο $\sum_{i=1}^{N} X_i = m+1$ το διάστημα $(m+\frac{1}{2}, m+\frac{3}{2})$, κ.ο.κ., και τελικά στο ενδεχόμενο $\sum_{i=1}^{N} X_i = M$ το διάστημα $(M-\frac{1}{2}, M+\frac{1}{2})$. Παρατηρήστε μάλιστα πως όταν $M = m$ η πρώτη προσέγγιση θα δώσει πιθανότητα ακριβώς 0. Δείτε τα ακόλουθα παραδείγματα.

Παράδειγμα 10.15. Για $N = 150$ ρίψεις ενός δίκαιου νομίσματος ($p = 1/2$), ποια είναι η πιθανότητα οι συνολικές κορώνες να είναι μεταξύ των 90 και των 105;

Για να απαντήσουμε, έστω $X_1, X_2, \ldots, X_N$ Τ.Μ. όπου $X_i = 1$ αν η i-οστή ρίψη είναι κορώνα, και $X_i = 0$ αλλιώς. Η μέση τιμή των X_i είναι $\frac{1}{2}$ και η διασπορά $p(1-p) = \frac{1}{4}$, κατά τα γνωστά για την κατανομή Bernoulli. Έστω επίσης $Y = \sum_{i=1}^{N} X_i$ το άθροισμα από κορώνες. Παρατηρούμε πως

$$\begin{aligned}
&P(90 \leq Y \leq 105)\\
&= P\left(90 - \frac{1}{2} \leq Y \leq 105 + \frac{1}{2}\right)\\
&= P\left(90 - \frac{1}{2} - 150 \times 0.5 \leq Y - 150 \times 0.5 \leq 105 + \frac{1}{2} - 150 \times 0.5\right)\\
&= P\left(\frac{90 - \frac{1}{2} - 150 \times 0.5}{\sqrt{150\frac{1}{4}}} \leq \frac{Y - 150 \times 0.5}{\sqrt{150\frac{1}{4}}} \leq \frac{105 + \frac{1}{2} - 150 \times 0.5}{\sqrt{150\frac{1}{4}}}\right)\\
&\simeq \Phi(4.98) - \Phi(2.37) \simeq 1 - 0.9911 = 0.0089,
\end{aligned}$$

Αν δεν είχαμε τροποποιήσει αρχικά το εύρος τιμών, το αποτέλεσμα θα ήταν

$$\Phi(4.90) - \Phi(2.45) \simeq 1 - 0.9928 = 0.0072,$$

Η ακριβής τιμή, όπως προκύπτει με χρήση υπολογιστή, είναι 0.0088. Η τροποποιημένη μεθοδολογία δίνει καλύτερο αποτέλεσμα, αλλά πάλι υπάρχει μια μικρή διαφορά, που οφείλεται στο ότι η ζητούμενη πιθανότητα είναι πολύ μικρή.

Παράδειγμα 10.16. *(Τζένγκις Χαν)* Νεώτερες αρχαιολογικές και ιστορικές έρευνες επιβεβαίωσαν την παραδοσιακή δοξασία ότι ο Τζένγκις Χαν είχε ακριβώς 10^4 παιδιά (γιους και κόρες). Υποθέτοντας ότι το φύλλο του καθενός παιδιού είναι ανεξάρτητο από τα άλλα, και ότι το κάθε παιδί είναι γιος με πιθανότητα 50%, θα υπολογίσουμε προσεγγιστικά, με χρήση του Κ.Ο.Θ., την πιθανότητα οι γιοι του Τζένγκις Χαν να ήταν στο πλήθος μεταξύ 4900 και 5100 (συμπεριλαμβανομένων των άκρων 4900, 5100).

Έστω X_i, $i = 1, \ldots, 10000$, τυχαίες μεταβλητές Bernoulli. Η $X_i = 1$ αν το i-οστό παιδί είναι αγόρι, και $X_i = 0$ αν το i-οστό παιδί είναι κορίτσι. Οι X_i έχουν μέση τιμή $\mu = \frac{1}{2}$ και διασπορά $\sigma^2 = \frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{4}$, συνεπώς τυπική απόκλιση $\sigma = \frac{1}{2}$. Για την ζητούμενη πιθανότητα έχουμε:

$$
\begin{aligned}
& P(4900 \leq \sum_{i=1}^{10000} X_i \leq 5100) \\
&= P(4899.5 \leq \sum_{i=1}^{10000} X_i \leq 5100.5) \\
&= P(4899.5 - 10000 \times \frac{1}{2} \leq \sum_{i=1}^{10000} X_i - 10000 \times \frac{1}{2} \leq 5100.5 - 10000 \times \frac{1}{2}) \\
&= P\left(\frac{4899.5 - 10000 \times \frac{1}{2}}{\frac{1}{2}\sqrt{10000}} \leq \frac{\sum_{i=1}^{10000} X_i - 10000 \times \frac{1}{2}}{\frac{1}{2}\sqrt{10000}} \leq \frac{5100.5 - 10000 \times \frac{1}{2}}{\frac{1}{2}\sqrt{10000}}\right) \\
&= P\left(-2.01 \leq \frac{\sum_{i=1}^{10000} X_i - 10000 \times \frac{1}{2}}{\frac{1}{2}\sqrt{10000}} \leq 2.01\right) \\
&\simeq P\left(-2.01 \leq Z \leq 2.01\right) = \Phi(2.01) - \Phi(-2.01) = 2\Phi(2.01) - 1 \simeq 0.955574.
\end{aligned}
$$

Χωρίς την τροποποίηση της πρώτης ισότητας, το αποτέλεσμα θα προέκυπτε $\simeq 0.954499$. Η ακριβής τιμή, με χρήση υπολογιστή, είναι $\simeq 0.955574$.

Παράδειγμα 10.17. Έστω πως η διάρκεια κάθε τηλεφωνικής κλήσης σε ένα δίκτυο έχει εκθετική κατανομή με μέση διάρκεια 85 δευτερόλεπτα, και έστω πως οι διάρκειες των διαδοχικών κλήσεων είναι ανεξάρτητες. Μια κλήση θεωρείται «σύντομη» αν διαρκέσει λιγότερο από ένα λεπτό. Ποια είναι η πιθανότητα, το πλήθος των σύντομων κλήσεων ανάμεσα στις 250 που έγιναν σε μία μέρα να είναι μικρότερο 120;

Έστω μια Τ.Μ. $Y \sim \text{Εκθ}(85)$. Η πιθανότητα μια κλήση να είναι σύντομη είναι

$$p = P(Y \leq 60) = 1 - e^{-60/85} \simeq 0.5063.$$

Έστω, τώρα, ανεξάρτητες Bern(p) Τ.Μ. X_i, όπου η κάθε $X_i = 1$ αν η κλήση i είναι σύντομη. Η μέση τιμή των X_i είναι $E(X_i) = p$ και η διασπορά τους $\sigma^2 = p(1-p)$. Η ζητούμενη πιθανότητα είναι

$$
\begin{aligned}
P\left(\sum_{i=1}^{250} X_i < 120\right) &= P\left(\sum_{i=1}^{250} X_i \leq 119.5\right) = P\left(\sum_{i=1}^{250}(X_i - p) \leq 119.5 - 250p\right) \\
&= P\left(\frac{1}{\sqrt{250p(1-p)}} \leq \sum_{i=1}^{250}(X_i - \mu) \leq \frac{119.5 - 250p}{\sqrt{250p(1-p)}}\right) \\
&\simeq P(Z \leq -0.8959) = \Phi(-0.8959) \simeq 0.1851,
\end{aligned}
$$

Αν δεν είχαμε τροποποιήσει αρχικά το εύρος τιμών, το αποτέλεσμα θα ήταν (επιβεβαιώστε το) $\Phi(-0.8326) \simeq 0.2025$. Η ακριβής τιμή, όπως προκύπτει με χρήση υπολογιστή, είναι επίσης $\simeq 0.1851$! Για την ακρίβεια, το σφάλμα του πρώτου μας υπολογισμού είναι περίπου ίσο με 10^{-6}.

Παρατήρηση: Παρατηρήστε ότι στα άνω δύο παραδείγματα είχαμε την προσέγγιση ενός αθροίσματος όμοια κατανεμημένων και ανεξάρτητων Τ.Μ. Bernoulli μέσω της κανονικής κατανομής. Παρατηρήστε όπως ότι ένα άθροισμα όμοια κατανεμημένων και ανεξάρτητων Τ.Μ. Bernoulli ακολουθεί την διωνυμική κατανομή. Άρα, μπορούμε να προσεγγίσουμε την διωνυμική κατανομή μέσω της κανονικής! Αυστηρά, έχουμε το ακόλουθο λήμμα, που δίνουμε χωρίς απόδειξη:

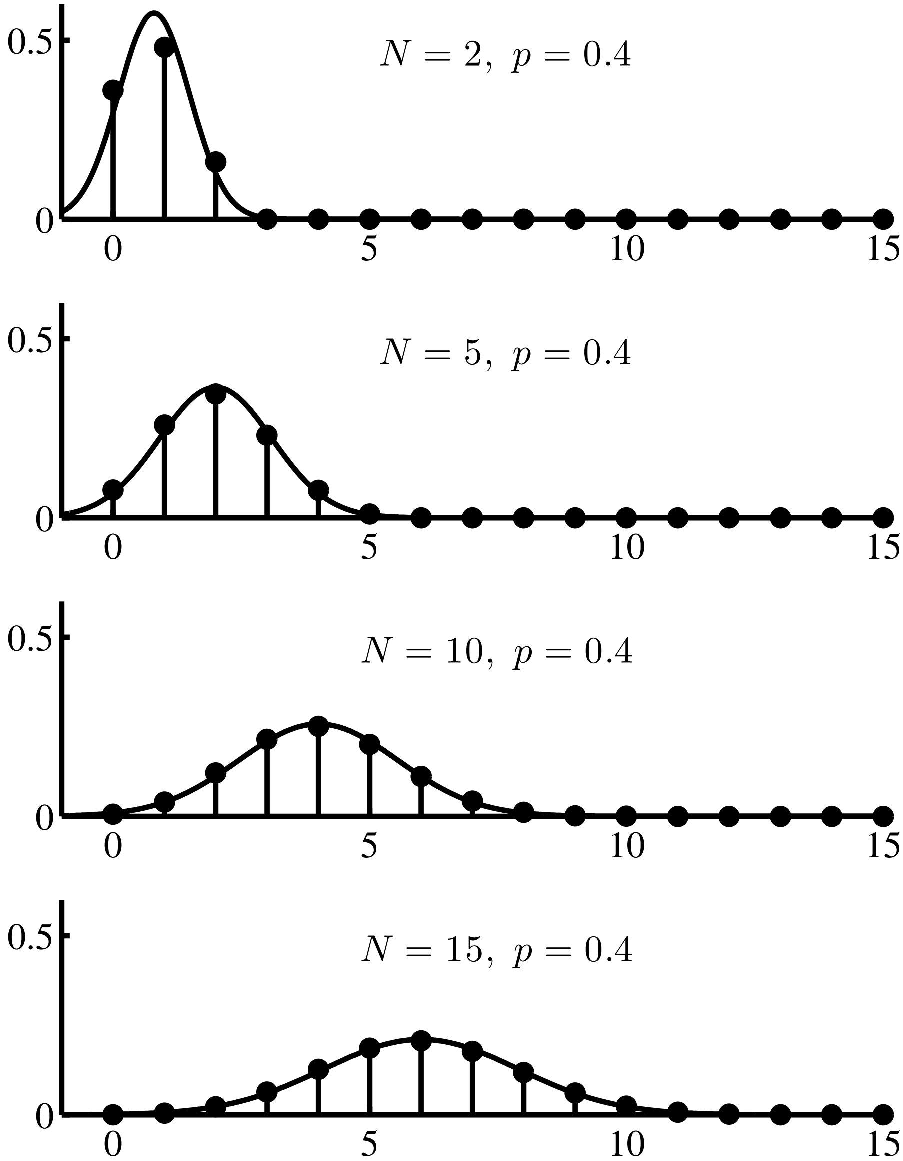
Λήμμα 10.3. (Προσέγγιση της Διωνυμικής Κατανομής από την Κανονική) *Έστω ακολουθία από διωνυμικές Τ.Μ. X_N με παραμέτρους N, p. Τότε, καθώς $N \to \infty$:*

$$P\left(a \le \frac{X_N - Np}{\sqrt{Np(1-p)}} \le b\right) \;\to\; \Phi(b) - \Phi(a),$$
$$P\left(\frac{X_N - Np}{\sqrt{Np(1-p)}} \le b\right) \;\to\; \Phi(b),$$
$$P\left(\frac{X_N - Np}{\sqrt{Np(1-p)}} \ge a\right) \;\to\; 1 - \Phi(a).$$

Παρατηρήσεις

1. Στο Σχήμα 10.2 έχουμε σχεδιάσει μια ακολουθία από διωνυμικές μάζες Τ.Μ. με p και $N = 2, 5, 10, 15$. Παρατηρήστε πως, όπως προβλέπει το Κ.Ο.Θ., καθώς μεγαλώνει το N, οι μάζες τείνουν στην κανονική κατανομή

2. Πρακτικά, η άνω προσέγγιση είναι ικανοποιητική όταν το N είναι αρκετά μεγάλο, π.χ., μεγαλύτερο του 50, και ταυτόχρονα το Np είναι αρκετά μεγαλύτερο της μονάδας, π.χ. μεγαλύτερο του 10.



Σχήμα 10.2: Σύγκλιση της διωνυμικής κατανομής στην κανονική, για $p = 0.4$ και για αυξανόμενες τιμές του N ($N = 2, 5, 10, 15$.)

Βιβλιογραφία

1. *Εισαγωγή στη Θεωρία Πιθανοτήτων και Εφαρμογές*, Μ. Κούτρας, Εκδόσεις Σταμούλη, 2012.
2. *Βασικές Αρχές Θεωρίας Πιθανοτήτων*, S. Ross, Κλειδάριθμος, 2012.
3. *Εισαγωγή στις Πιθανότητες*, Δ. Μπερτσεκάς, Ι. Τσιτσικλής, Εκδόσεις Τζιόλα, 2010.
4. *Εισαγωγή στη Θεωρία Πιθανοτήτων*, P. G. Hoel, S. C. Port, C. J. Stone, Πανεπιστημιακές Εκδόσεις Κρήτης, 2011.
5. *A First Course in Probability*, S. Ross, Prentice Hall, 2009.
6. *Introduction to Probability*, D. P. Bertsekas, J. N. Tsitsiklis, Athena Scientific, 2008.
7. *Introduction to Probability Theory*, P. G. Hoel, S. C. Port, C. J. Stone, Brooks Cole, 1972.

Ευρετήριο